

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 1, 2004, 14:52:43 ; Search time 53.2407 Seconds
(without alignments)
1302.928 Million cell updates/sec

Title: US-10-620-039-1_COPY_1_125

Perfect score: 125
Sequence: 1 TTGGCCCACTCCCTCTCTGCG.....CGCAGAGAGGAGTGGCCAA 125

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	125	100.0	145	1	US-07-789-917A-1
2	125	100.0	145	3	US-08-702-573-4
3	125	100.0	145	3	US-08-525-866-1
4	125	100.0	145	3	US-07-982-193-1
5	125	100.0	165	1	US-07-989-841A-1
6	125	100.0	165	2	US-08-440-738A-1
7	125	100.0	165	3	US-08-471-914-1
8	125	100.0	165	4	US-09-276-625-7
9	125	100.0	192	3	US-08-702-573-3
10	125	100.0	4680	1	US-08-254-358-1
11	125	100.0	4680	1	US-08-475-391-1
12	125	100.0	4680	2	US-08-709-609-1
13	125	100.0	4680	5	PCT-US95-07178-1
14	125	100.0	5932	4	US-09-295-141-4
15	125	100.0	5932	4	US-09-299-141-4
16	125	100.0	6142	4	US-09-299-141-8
17	125	100.0	6142	4	US-09-299-141-8
18	125	100.0	6253	3	US-08-893-327-15
19	125	100.0	6253	3	US-08-893-327-15
20	125	100.0	6280	3	US-08-893-327-17
21	125	100.0	6280	3	US-08-893-327-17
22	125	100.0	6280	3	US-08-893-327-19
23	125	100.0	6280	3	US-08-893-327-19
24	125	100.0	6565	4	US-09-299-141-1
25	125	100.0	6565	4	US-09-299-141-1
26	125	100.0	6714	4	US-09-299-141-6
27	125	100.0	6714	4	US-09-299-141-6

28	125	100.0	6924	4	US-09-299-141-9	Sequence 9, Appli
29	125	100.0	6924	4	US-09-299-141-9	Sequence 9, Appli
30	125	100.0	6924	4	US-09-299-141-10	Sequence 10, Appl
31	125	100.0	6924	4	US-09-299-141-10	Sequence 10, Appl
32	125	100.0	6924	4	US-09-299-141-11	Sequence 11, Appl
33	125	100.0	6924	4	US-09-299-141-11	Sequence 11, Appl
34	125	100.0	6981	4	US-09-299-141-7	Sequence 7, Appli
35	125	100.0	6981	4	US-09-299-141-7	Sequence 7, Appli
36	125	100.0	7054	4	US-09-299-141-3	Sequence 3, Appli
37	125	100.0	7054	4	US-09-299-141-3	Sequence 3, Appli
38	125	100.0	7405	4	US-09-299-141-2	Sequence 2, Appli
39	125	100.0	7405	4	US-09-299-141-2	Sequence 2, Appli
40	125	100.0	7492	4	US-09-299-141-5	Sequence 5, Appli
41	125	100.0	7492	4	US-09-299-141-5	Sequence 5, Appli
42	125	100.0	8698	4	US-09-770-315-2	Sequence 4, Appli
43	123.4	98.7	272	4	US-09-276-625-4	Sequence 4, Appli
44	123	98.4	174	4	US-09-394-110A-1	Sequence 1, Appli
45	121.8	97.4	5585	2	US-08-305-221-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

```

US-07-789-917A-1
; Sequence 1, Application US/07789917A
; Patent No. 5252479
; GENERAL INFORMATION:
; APPLICANT: Srivastava, Arun
; TITLE OF INVENTION: SAFE VECTOR FOR GENE THERAPY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Scully, Scott, Murphy Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release q.0, Version q.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07789,917A
; FILING DATE: 19911118
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McNulty, William E.
; REGISTRATION NUMBER: 22,606
; REFERENCE/DOCKET NUMBER: 8361
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-789-917A-1

```

Query Match 100.0%; Score 125; DB 1; Length 145;

Best Local Similarity 100.0%; Pred. No. 7e-25;

Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTGGCCCACTCCCTCTCTGCGCGCTCGCTGAGCGCGGCGGACCAAGGTGCGC 60

Db 1 TTGGCCCACTCCCTCTCTGCGCGCTCGCTGAGCGCGGCGGACCAAGGTGCGC 60

Qy 61 CGACGCCCGCGGCTTTGCCCGCGCGCTTCAGTGAGCGCGCGGAGGAGTG 120

Db 61 CGAGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGAGCGCGAGAGGAGTG 120
Qy 121 GCCAA 125
Db 121 GCCAA 125

RESULT 2
US-08-702-573-4
; Sequence 4, Application US/08702573
; Patent No. 6033885
; GENERAL INFORMATION:
; APPLICANT: LATTI, Martine
; APPLICANT: DENEFE, Patrice
; APPLICANT: VIGNE, Emmanuelle
; APPLICANT: FERRICAUD, Michel
; TITLE OF INVENTION: INTEGRATIVE RECOMBINANT ADENOVIRUSES,
; TITLE OF INVENTION: PREPARATION THEREOF AND THERAPEUTICAL USES THEREOF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd. 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,573
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94/02445
; FILING DATE: 03-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR95/00233
; FILING DATE: 28-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: 38,619
; REFERENCE/DOCKET NUMBER: ST94011-US
; TELEPHONE: (610)454-3839
; TELEFAX: (610)454-3808
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..145
; OTHER INFORMATION: /note="Minimal ITR Sequence"
US-08-702-573-4

Query Match 100.0%; Score 125; DB 3; Length 145;
Best Local Similarity 100.0%; Pred. No. 7e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTGGCCACTCCCTCTCTCGGGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
Db 1 TTGGCCACTCCCTCTCTCGGGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
Qy 61 CGAGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGAGCGCGAGAGGAGTG 120
Db 61 CGAGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGAGTG 120

Qy 121 GCCAA 125
Db 121 GCCAA 125
RESULT 3
US-08-525-866-1/c
; Sequence 1, Application US/08525866
; Patent No. 6207457
; GENERAL INFORMATION:
; APPLICANT: NATSOULIS, GEORGES
; APPLICANT: FUROSKY, RICHARD T.
; TITLE OF INVENTION: TARGETED NUCLEOTIDE SEQUENCE DELIVERY
; TITLE OF INVENTION: AND INTEGRATION SYSTEM
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: REED & ROBINS
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/525,866
; FILING DATE: 08-SEP-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 0800-0006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 327-3400
; TELEFAX: (415) 327-3231
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-525-866-1

Query Match 100.0%; Score 125; DB 3; Length 145;
Best Local Similarity 100.0%; Pred. No. 7e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTGGCCACTCCCTCTCTCGGGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
Db 125 TTGGCCACTCCCTCTCTCGGGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 66
Qy 61 CGAGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGAGCGCGAGAGGAGTG 120
Db 65 CGAGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGAGTG 6
Qy 121 GCCAA 125
Db 5 GCCAA 1

RESULT 4
US-07-982-193-1
; Sequence 1, Application US/07982193
; Patent No. 6261834
; GENERAL INFORMATION:
; APPLICANT: Srivastava, Arun
; TITLE OF INVENTION: SAFE VECTOR FOR GENE THERAPY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:

```

; ADDRESSES: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/982,193
; FILING DATE: 19921125
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McNulty, William E.
; REGISTRATION NUMBER: 22,606
; REFERENCE/DOCKET NUMBER: 8361
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-07-982-193-1

Query Match 100.0%; Score 125; DB 3; Length 145;
Best Local Similarity 100.0%; Pred. No. 7e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCC 60
Db 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCC 60
QY 61 CGAGCGCGCGGCTTTCGCGCGGCGCTCACTGAGCGCGGCGACGAGAGGGAGTG 120
Db 61 CGAGCGCGCGGCTTTCGCGCGGCGCTCACTGAGCGCGGCGACGAGAGGGAGTG 120
QY 121 GCCAA 125
Db 121 GCCAA 125

RESULT 5
US-07-989-841A-1
; Sequence 1, Application US/07989841A
; Patent No. 5478745
; GENERAL INFORMATION:
; APPLICANT: Samulski, R. J.
; APPLICANT: Xiao, X.
; TITLE OF INVENTION: Recombinant Viral Vector System
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/989,841A
; FILING DATE: On even date herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 6636-022
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 165 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; MOLECULE TYPE: DNA (genomic)
; US-07-989-841A-1

Query Match 100.0%; Score 125; DB 1; Length 165;
Best Local Similarity 100.0%; Pred. No. 7e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCC 60
Db 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCC 80
QY 61 CGAGCGCGCGGCTTTCGCGCGGCGCTCACTGAGCGCGGCGACGAGAGGGAGTG 120
Db 61 CGAGCGCGCGGCTTTCGCGCGGCGCTCACTGAGCGCGGCGACGAGAGGGAGTG 140
QY 121 GCCAA 125
Db 141 GCCAA 145

RESULT 6
US-08-440-738A-1
; Sequence 1, Application US/08440738A
; Patent No. 5869305
; GENERAL INFORMATION:
; APPLICANT: Samulski, R. J.
; APPLICANT: Xiao, X.
; TITLE OF INVENTION: Recombinant Viral Vector System
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,738A
; FILING DATE: May 15, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 6636-022
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 165 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

```

```
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-08-440-738A-1

Query Match      100.0%; Score 125; DB 2; Length 165;
Best Local Similarity 100.0%; Pred. No. 7e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCAGTGGAGCGAGCGAGAGGAGTG 120
DB 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCAGTGGAGCGAGCGAGAGGAGTG 80

QY 61 CGAGCGCCGGGCTTTGCGCGCGGCTCAGTGGAGCGAGCGAGAGGAGTG 120
DB 81 CGAGCGCCGGGCTTTGCGCGCGGCTCAGTGGAGCGAGCGAGAGGAGTG 140

QY 121 GCCAA 125
DB 141 GCCAA 145

RESULT 7
US-08-471-914-1
; Sequence 1, Application US/08471914A
; Patent No. 6057152
; GENERAL INFORMATION:
; APPLICANT: Samulski, R.
; APPLICANT: Xiao, X.
; TITLE OF INVENTION: RECOMBINANT VIRAL VECTOR SYSTEM
; FILE REFERENCE: 6636-027
; CURRENT APPLICATION NUMBER: US/08/471.914A
; CURRENT FILING DATE: 1995-06-06
; EARLIER APPLICATION NUMBER: 08/440.738
; EARLIER FILING DATE: 1995-05-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: double-D
US-08-471-914-1

Query Match      100.0%; Score 125; DB 3; Length 165;
Best Local Similarity 100.0%; Pred. No. 7e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCAGTGGAGCGAGCGAGAGGAGTG 60
DB 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCAGTGGAGCGAGCGAGAGGAGTG 80

QY 61 CGAGCGCCGGGCTTTGCGCGCGGCTCAGTGGAGCGAGCGAGAGGAGTG 120
DB 81 CGAGCGCCGGGCTTTGCGCGCGGCTCAGTGGAGCGAGCGAGAGGAGTG 140

QY 121 GCCAA 125
DB 141 GCCAA 145

RESULT 8
US-09-276-625-7
; Sequence 7, Application US/09276625
; Patent No. 6436392
; GENERAL INFORMATION:
; APPLICANT: Engelhardt, John F.
; APPLICANT: Duan, Dongsheng
; TITLE OF INVENTION: Adeno-associated virus vectors
; FILE REFERENCE: 875.007US1
; CURRENT APPLICATION NUMBER: US/09/276.625
; CURRENT FILING DATE: 1999-03-25
```

```
; PRIOR APPLICATION NUMBER: US 60/086,166
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: SEQ ID NO:1 of U.S. Patent No. 6436392 5,478,745
US-09-276-625-7

Query Match      100.0%; Score 125; DB 4; Length 165;
Best Local Similarity 100.0%; Pred. No. 7e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCAGTGGAGCGAGCGAGAGGAGTG 60
DB 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCAGTGGAGCGAGCGAGAGGAGTG 80

QY 61 CGAGCGCCGGGCTTTGCGCGCGGCTCAGTGGAGCGAGCGAGAGGAGTG 120
DB 81 CGAGCGCCGGGCTTTGCGCGCGGCTCAGTGGAGCGAGCGAGAGGAGTG 140

QY 121 GCCAA 125
DB 141 GCCAA 145

RESULT 9
US-08-702-573-3
; Sequence 3, Application US/08702573
; Patent No. 6033885
; GENERAL INFORMATION:
; APPLICANT: Latta, Martine
; APPLICANT: DENEZLE, Patrice
; APPLICANT: VIGNE, Emmanuelle
; APPLICANT: FERRICAUDET, Michel
; TITLE OF INVENTION: INTEGRATIVE RECOMBINANT ADENOVIRUSES,
; NUMBER OF SEQUENCES: 13
; PREPARATION THEREOF AND THERAPEUTICAL USES THEREOF
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd. 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702.573
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94/02445
; FILING DATE: 03-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR95/00233
; FILING DATE: 28-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: 38,619
; REFERENCE/DOCKET NUMBER: ST94011-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)454-3839
; TELEFAX: (610)454-3808
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 192 base pairs
```


Query Match 100.0%; Score 125; DB 1; Length 4680;
Best Local Similarity 100.0%; Pred. No. 7.9e-25;

	Query Match	100.0%;	Score 125;	DB 1;	Length 4680;
	Best Local Similarity	100.0%;	Pred. NO. 7.9e-25;		
	Matches 125;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	TTGGCCACTCCCTCTCTGGCGGCTGCCTCGCTCACTGAGCGCGGGCCACCAAGGTGCC	60		
Db	1	TTGGCCACTCCCTCTCTGGCGGCTGCCTCGCTCACTGAGCGCGGGCCACCAAGGTGCC	60		
Qy	61	CGAGCGCCGGCGCTTTGGCCGGGGCGGCTCAGTGAGCGAGCGCGCCAGAGGGGAGTG	120		
Db	61	CGAGCGCCGGCGCTTTGGCCGGGGCGGCTCAGTGAGCGAGCGCGCCAGAGGGGAGTG	120		
Qy	121	GCCAA 125			

```
Db      121 GCCAA 125
|||||
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: PCT/US95/07178
  FILING DATE:
  CLASSIFICATION:
  ATTORNEY/AGENT INFORMATION:
    NAME: Noland, Greta E.
    REGISTRATION NUMBER: 35,302
    REFERENCE/DOCKET NUMBER: 31975
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (312) 474-6300
    TELEFAX: (312) 474-0448
    TELEX: 25-3856
  INFORMATION FOR SEQ ID NO: 1:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 4680 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
PCT-US95-07178-1
Query Match      100.0%; Score 125; DB 2; Length 4680;
Best Local Similarity 100.0%; Pred. No. 7.9e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGGCGGACCAAGGTCGCC 60
Db      1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGGCGGACCAAGGTCGCC 60

QY      61 CGAGCGCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGGCGGAGGAGTG 120
Db      61 CGAGCGCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGGCGGAGGAGTG 120

QY      121 GCCAA 125
Db      121 GCCAA 125

RESULT 14
US-09-299-141-4
; Sequence 4, Application US/09299141
; Patent No. 6461606
; GENERAL INFORMATION:
; APPLICANT: FLOTTE, TERENCE R.
; APPLICANT: SONG, SIHONG
; APPLICANT: BYRNE, BARRY J.
; APPLICANT: MORGAN, MICHAEL
; TITLE OF INVENTION: MATERIALS AND METHODS FOR GENE THERAPY
; FILE REFERENCE: 4300.011800
; CURRENT APPLICATION NUMBER: US/09/299,141
; CURRENT FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: 60/083,025
; EARLIER FILING DATE: 1998-04-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 5932
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:p43C-AT
US-09-299-141-4

Db      121 GCCAA 125
|||||
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/709,609
  FILING DATE:
  CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
    NAME: No 585877sand, Greta E.
    REGISTRATION NUMBER: 35,302
    REFERENCE/DOCKET NUMBER: 31975
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (312) 474-6300
    TELEFAX: (312) 474-0448
    TELEX: 25-3856
  INFORMATION FOR SEQ ID NO: 1:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 4680 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
US-08-709-609-1
Query Match      100.0%; Score 125; DB 2; Length 4680;
Best Local Similarity 100.0%; Pred. No. 7.9e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGGCGGACCAAGGTCGCC 60
Db      1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGGCGGACCAAGGTCGCC 60

QY      61 CGAGCGCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGGCGGAGGAGTG 120
Db      61 CGAGCGCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGGCGGAGGAGTG 120

QY      121 GCCAA 125
Db      121 GCCAA 125

RESULT 13
PCT-US95-07178-1
; Sequence 1, Application PC/TUS9507178
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
  ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
  STREET: 6300 Sears Tower, 233 S. Wacker Drive
  CITY: Chicago
  STATE: Illinois
  COUNTRY: USA
  ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/709,609
  FILING DATE:
  CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
    NAME: No 585877sand, Greta E.
    REGISTRATION NUMBER: 35,302
    REFERENCE/DOCKET NUMBER: 31975
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (312) 474-6300
    TELEFAX: (312) 474-0448
    TELEX: 25-3856
  INFORMATION FOR SEQ ID NO: 1:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 4680 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
US-08-709-609-1
Query Match      100.0%; Score 125; DB 2; Length 4680;
Best Local Similarity 100.0%; Pred. No. 7.9e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGGCGGACCAAGGTCGCC 60
Db      1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGGCGGACCAAGGTCGCC 60

QY      61 CGAGCGCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGGCGGAGGAGTG 120
Db      61 CGAGCGCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGGCGGAGGAGTG 120

QY      121 GCCAA 125
Db      121 GCCAA 125

RESULT 13
PCT-US95-07178-1
; Sequence 1, Application PC/TUS9507178
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
  ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
  STREET: 6300 Sears Tower, 233 S. Wacker Drive
  CITY: Chicago
  STATE: Illinois
  COUNTRY: USA
  ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: PCT/US95/07178
  FILING DATE:
  CLASSIFICATION:
  ATTORNEY/AGENT INFORMATION:
    NAME: Noland, Greta E.
    REGISTRATION NUMBER: 35,302
    REFERENCE/DOCKET NUMBER: 31975
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (312) 474-6300
    TELEFAX: (312) 474-0448
    TELEX: 25-3856
  INFORMATION FOR SEQ ID NO: 1:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 4680 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
PCT-US95-07178-1
Query Match      100.0%; Score 125; DB 5; Length 4680;
Best Local Similarity 100.0%; Pred. No. 7.9e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGGCGGACCAAGGTCGCC 60
Db      1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGGCGGACCAAGGTCGCC 60

QY      61 CGAGCGCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGGCGGAGGAGTG 120
Db      61 CGAGCGCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGGCGGAGGAGTG 120

QY      121 GCCAA 125
Db      121 GCCAA 125

RESULT 14
US-09-299-141-4
; Sequence 4, Application US/09299141
; Patent No. 6461606
; GENERAL INFORMATION:
; APPLICANT: FLOTTE, TERENCE R.
; APPLICANT: SONG, SIHONG
; APPLICANT: BYRNE, BARRY J.
; APPLICANT: MORGAN, MICHAEL
; TITLE OF INVENTION: MATERIALS AND METHODS FOR GENE THERAPY
; FILE REFERENCE: 4300.011800
; CURRENT APPLICATION NUMBER: US/09/299,141
; CURRENT FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: 60/083,025
; EARLIER FILING DATE: 1998-04-24
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 5932
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:p43C-AT
US-09-299-141-4
```

Query Match 100.0%; Score 125; DB 4; Length 5932;
 Best Local Similarity 100.0%; Pred. No. 8e-25;
 Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCC 60
 Db 18 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCC 77

Qy 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGAGTG 120
 Db 78 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGAGTG 137

Qy 121 GCCAA 125
 Db 138 GCCAA 142

RESULT 15
 US-09-299-141-4/c
 ; Sequence 4, Application US/09299141
 ; Patent No. 6461606
 ; GENERAL INFORMATION:
 ; APPLICANT: FLOTTE, TERENCE R.
 ; APPLICANT: SONG, SIHONG
 ; APPLICANT: BYRNE, BARRY J.
 ; APPLICANT: MORGAN, MICHAEL
 ; TITLE OF INVENTION: MATERIALS AND METHODS FOR GENE THERAPY
 ; FILE REFERENCE: 4300.011800
 ; CURRENT APPLICATION NUMBER: US/09/299,141
 ; CURRENT FILING DATE: 1999-04-23
 ; EARLIER APPLICATION NUMBER: 60/083,025
 ; EARLIER FILING DATE: 1998-04-24
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 4
 ; LENGTH: 5932
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:p43C-AT
 US-09-299-141-4

Query Match 100.0%; Score 125; DB 4; Length 5932;
 Best Local Similarity 100.0%; Pred. No. 8e-25;
 Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCC 60
 Db 3078 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCC 3019

Qy 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGAGTG 120
 Db 3018 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGAGTG 2959

Qy 121 GCCAA 125
 Db 2958 GCCAA 2954

Search completed: April 1, 2004, 17:04:17
 Job time : 53.2407 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 1, 2004, 14:05:53 ; Search time 1573.61 Seconds
(without alignments) 3442.959 Million cell updates/sec

Title: US-10-620-039-1_COPY_1_125
Perfect score: 125
Sequence: 1 TTGGCCACTCCTCTCTGCG.....CGCAGAGGAGGAGGGCAA 125

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenBank:
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_ov.*
5: gb_ov.*
6: gb_pat.*
7: gb_pat.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vt.*
15: em_ba.*
16: em_fun.*
17: em_fun.*
18: em_in.*
19: em_mu.*
20: em_mu.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vt.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	125	100.0	130	6	AX384518	Sequence
2	125	100.0	145	6	A46401	Sequence 4
3	125	100.0	145	6	AR140333	Sequence
4	125	100.0	145	6	AX286293	Sequence
5	125	100.0	145	14	AA2LTR1	K01624 Adeno-assoc
6	125	100.0	145	14	AA2LTR2	M10681 Adeno-assoc
7	125	100.0	145	14	AA2REPORI	AR034135 Sequence
8	125	100.0	165	6	AR034135	I16806 Sequence 1
9	125	100.0	165	6	I16806	AR223306 Sequence
10	125	100.0	185	6	AX106702	Sequence
11	125	100.0	185	6	BD18219	BD218219 Adeno-ass
12	125	100.0	192	6	A46400	Sequence 3
13	125	100.0	198	6	AX418199	Sequence
14	125	100.0	198	6	AX703496	AX703496 Sequence
15	125	100.0	207	6	AX703496	J01902 adeno-assoc
16	125	100.0	2116	14	AA2LEFT	AX135805 Sequence
17	125	100.0	4675	6	AX135805	AX286292 Sequence
18	125	100.0	4675	6	AX286292	AX73252 Sequence
19	125	100.0	4675	6	AX73252	BD094552 Method of
20	125	100.0	4675	6	BD094552	J01901 Adeno-assoc
21	125	100.0	4675	14	AA2CG	AX282480 Sequence
22	125	100.0	4679	6	AX282480	AF043303 Adeno-ass
23	125	100.0	4679	14	AF043303	AR028767 Sequence
24	125	100.0	4680	6	AR028767	162303 Sequence 1
25	125	100.0	4680	6	T62303	BD242774 Adeno-ass
26	125	100.0	4681	6	BD242774	BD242775 Adeno-ass
27	125	100.0	4683	6	BD242775	AF028704 Adeno-ass
28	125	100.0	4683	14	AF028704	AR235457 Sequence
29	125	100.0	5932	6	AR235457	AR235461 Sequence
30	125	100.0	5932	6	AR235457	AR235461 Sequence
31	125	100.0	6142	6	AR235461	AR235454 Sequence
32	125	100.0	6142	6	AR235461	AR235459 Sequence
33	125	100.0	6565	6	AR235454	AR235459 Sequence
34	125	100.0	6565	6	AR235454	AR235462 Sequence
35	125	100.0	6714	6	AR235459	AR235463 Sequence
36	125	100.0	6714	6	AR235459	AR235463 Sequence
37	125	100.0	6924	6	AR235462	AR235464 Sequence
38	125	100.0	6924	6	AR235462	AR235460 Sequence
39	125	100.0	6924	6	AR235463	AR235460 Sequence
40	125	100.0	6924	6	AR235463	AR235456 Sequence
41	125	100.0	6924	6	AR235464	AR235456 Sequence
42	125	100.0	6924	6	AR235464	AR235456 Sequence
43	125	100.0	6981	6	AR235460	AR235456 Sequence
44	125	100.0	6981	6	AR235460	AR235456 Sequence
45	125	100.0	7054	6	AR235456	AR235456 Sequence

ALIGNMENTS

RESULT 1	AX384518	AX384518	Sequence 1 from Patent WO0214526.	130 bp	DNA	Linear	PAT 19-MAR-2002
LOCUS	AX384518	Sequence 1 from Patent WO0214526.					
DEFINITION	AX384518	Sequence 1 from Patent WO0214526.					
ACCESSION	AX384518	Sequence 1 from Patent WO0214526.					
VERSION	AX384518.1	GI:19577720					
KEYWORDS							
SOURCE		Adeno-associated virus 2					
ORGANISM		Adeno-associated virus 2					
REFERENCE		Li, S.					
AUTHORS		Replication competent aav helper functions					
TITLE		Patent: WO 0214526-A 1 21-FEB-2002;					
JOURNAL		Neurologix, Inc. (US)					

```
FEATURES
  source      Location/Qualifiers
1..130
  /organism="Adeno-associated virus 2"
  /mol_type="unassigned DNA"
  /db_xref="taxon:10804"
ORIGIN
Query Match      100.0%; Score 125; DB 6; Length 130;
Best Local Similarity 100.0%; Pred. No. 2.8e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
Db 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
Db 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
QY 121 GCCAA 125
Db 121 GCCAA 125
RESULT 2
A46401
LOCUS      A46401      145 bp      DNA      linear      PAT 07-MAR-1997
DEFINITION Sequence 4 from Patent WO9523867.
ACCESSION A46401
VERSION A46401.1 GI:2300602
KEYWORDS  unclassified
SOURCE      unclassified
ORGANISM    unclassified.
REFERENCE 1 (bases 1 to 145)
AUTHORS Denefle,P., Letta,M., Perricaudet,M. and Vigne,E.
TITLE INTEGRATIVE RECOMBINANT ADENOVIRUSES, PREPARATION THEREOF AND
THERAPEUTICAL USES THEREOF
JOURNAL Patent: WO 9523867-A 4 08-SEP-1995;
COMMENT RHONE POULENC RORER SA (FR)
Other publication AU 1852695 950918
Other publication FR 2716893 950908.
FEATURES
  source      Location/Qualifiers
1..145
  /organism="unclassified"
  /mol_type="unassigned DNA"
  /db_xref="taxon:32644"
ORIGIN
Query Match      100.0%; Score 125; DB 6; Length 145;
Best Local Similarity 100.0%; Pred. No. 2.8e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
Db 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
Db 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
QY 121 GCCAA 125
Db 121 GCCAA 125
RESULT 3
AR140333/c
LOCUS      AR140333      145 bp      DNA      linear      PAT 16-JUN-2001
DEFINITION Sequence 1 from patent US 6207457.
ACCESSION AR140333
VERSION AR140333.1 GI:14482829
KEYWORDS
```

```
SOURCE      Unknown.
ORGANISM      Unclassified.
REFERENCE 1 (bases 1 to 145)
AUTHORS Natsoulis,G. and Surosky,R.T.
TITLE Targeted nucleotide sequence delivery and integration system
JOURNAL Patent: US 6207457-A 1 27-MAR-2001;
FEATURES
  source      Location/Qualifiers
1..145
  /organism="unknown"
  /mol_type="unassigned DNA"
ORIGIN
Query Match      100.0%; Score 125; DB 6; Length 145;
Best Local Similarity 100.0%; Pred. No. 2.8e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
Db 125 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 66
QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
Db 65 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 6
QY 121 GCCAA 125
Db 5 GCCAA 1
RESULT 4
AX286293
LOCUS      AX286293      145 bp      DNA      linear      PAT 21-NOV-2001
DEFINITION Sequence 2 from Patent WO0180840.
ACCESSION AX286293
VERSION AX286293.1 GI:17048541
KEYWORDS  Adeno-associated virus 2
SOURCE      Adeno-associated virus 2
ORGANISM    Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
REFERENCE 1
AUTHORS Raj,K. and Beard,P.M.
TITLE Cytotoxic agents
JOURNAL Patent: WO 0180840-A 2 01-NOV-2001;
JOURNAL BTG INTERNATIONAL LIMITED (GB)
FEATURES
  source      Location/Qualifiers
1..145
  /organism="Adeno-associated virus 2"
  /mol_type="unassigned DNA"
  /db_xref="taxon:10804"
  misc_structure 1..145
  /notes="ITR"
  misc_feature 72
  /notes="Unpaired base"
  misc_feature 94
  /notes="Unpaired base"
ORIGIN
Query Match      100.0%; Score 125; DB 6; Length 145;
Best Local Similarity 100.0%; Pred. No. 2.8e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
Db 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 80
QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
Db 81 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 140
QY 121 GCCAA 125
Db 141 GCCAA 145
```

LOCUS	AA2LTR2	145 bp	DNA	linear	VRL 27-APR-1993
DEFINITION	Adeno-associated virus 2 right terminal sequence.				
ACCESSION	K01625				
VERSION	K01625.1	GI:209624			
KEYWORDS	replication; terminal repeat.				
SEGMENT	2 of 2				
SOURCE	Adeno-associated virus 2H				
ORGANISM	Adeno-associated virus 2H				
REFERENCE	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.				
AUTHORS	1 (bases 1 to 145)				
TITLE	Lusby,E., Fife,K.H. and Berns,K.I.				
	Nucleotide sequence of the inverted terminal repetition in				
	adeno-associated virus DNA				
JOURNAL	J. Virol. 34 (2), 402-409 (1980)				
MEDLINE	80185149				
PUBMED	6246271				
REFERENCE	2 (bases 1 to 145)				
AUTHORS	LeFebvre,R.B., Riva,S. and Berns,K.I.				
TITLE	Confirmation takes precedence over sequence in adeno-associated				
	virus DNA replication				
JOURNAL	Mol. Cell. Biol. 4 (7), 1416-1419 (1984)				
MEDLINE	85061247				
PUBMED	6504049				
COMMENT	Original source text: Adeno-associated virus 2H DNA (clone pSM620				
	[2]), from KB or HeLa cells.				
	The focus of both papers is the method of replication of the virus.				
	[1] notes that the initial tt is present only 30% of the time; it				
	is shortened to t in 50% of the population and missing altogether				
	in 15% of the population. There is further sequence heterogeneity				
	which can be explained by assuming that the terminal 125 bases,				
	which form an imperfect palindrome, are replaced by their inverted				
	complement during replication. [2] found that deletion of the 9				
	terminal bases on the right and the 113 terminal bases on the left				
	of AA2 genome did not stop DNA replication. Further deletion of				
	an 11-base symmetrical sequence (bases 89 to 99) in the right				
	terminal repetition inhibits DNA replication. Substitution of				
	either an 8-base (cagatctg) or 12-base (cgcgatccgcg) symmetrical				
	sequence unrelated to the original 11-base sequence restores DNA				
	replication. All of this can be explained by assuming that the 125				
	base palindrome mentioned above form a t-shaped secondary structure				
	which provides a primer for DNA polymerase during replication.				
FEATURES	Location/Qualifiers				
source	1. .145				
	/organism="Adeno-associated virus 2H"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:10805"				
ORIGIN	22 bases upstream of HaeIII sites.				
Query Match	100.0%; Score 125; DB 14; Length 145;				
Best Local Similarity	100.0%; Pred. No.2.8e-18;				
Matches 125; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
QY	1 TTGGCCACTCCCTCTCTGGCGGCTCGCTCGCTCACTGAGCGCGGGCGACCAAGGTGCC 60				
Db	21 TTGGCCACTCCCTCTCTGGCGGCTCGCTCGCTCACTGAGCGCGGGCGACCAAGGTGCC 80				
QY	61 CGACGCCCGGGCTTTTGGCCGGGGGGCGCTCTAGTGAGCGAGCGGCGCGCAGAGGGAGTG 120				
Db	81 CGACGCCCGGGCTTTTGGCCGGGGGGCGCTCTAGTGAGCGAGCGGCGCGCAGAGGGAGTG 140				
QY	121 GCCAA 125				
Db	141 GCCAA 145				
RESULT 7					
AA2REPORI/c					
LOCUS	AA2REPORI	145 bp ss-DNA	linear		VRL 27-APR-1993
DEFINITION	Adeno-associated virus origin of replication (genome 3' terminus).				
ACCESSION	M10681				
VERSION	M10681.1	GI:209626			
KEYWORDS					
SOURCE	Adeno-associated virus 2H				

```

ORGANISM
Adeno-associated virus 2H
Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
1 (bases 1 to 145)
AUTHORS
Berns, K.I., Hauswirth, W.W., Fife, K.H. and Luby, E.
TITLE
Adeno-associated virus DNA replication
JOURNAL
Cold Spring Harb. Symp. Quant. Biol. 43 Pt 2, 781-787 (1979)
MEDLINE
80023388
PUBMED
226321
COMMENT
Original source text: Adeno associated virus 2H (AAV2 H) DNA.
FEATURES
Location/Qualifiers
1..145
/organism="Adeno-associated virus 2H"
/mol_type="genomic DNA"
/db_xref="taxon:10805"

ORIGIN
Query Match 100.0%; Score 125; DB 14; Length 145;
Best Local Similarity 100.0%; Pred. No. 2.8e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 TTGGCCACTCCTCTCTCGGCGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTCGCC 60
DB 145 TTGGCCACTCCTCTCTCGGCGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTCGCC 86
QY 61 CGAGCGCCGGGCTTTGCCCGGCGGCGCTCACTAGCGCGGCGCGAGAGGGAGTG 120
DB 85 CGAGCGCCGGGCTTTGCCCGGCGGCGCTCACTAGCGAGCGAGCGCGAGAGGGAGTG 26
QY 121 GCCAA 125
DB 25 GCCAA 21

RESULT 8
AR034135
LOCUS
AR034135
DEFINITION
Sequence 1 from patent US 5869305.
ACCESSION
AR034135
VERSION
AR034135.1 GI:5949740
KEYWORDS
Unknown.
SOURCE
Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 165)
AUTHORS
Samulski, R.Jude. and Xiao, X.
TITLE
Recombinant viral vector system
JOURNAL
Patent: US 5869305-A 1 09-FEB-1999;
FEATURES
Location/Qualifiers
1..165
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 125; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.7e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 TTGGCCACTCCTCTCTCGGCGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTCGCC 60
DB 21 TTGGCCACTCCTCTCTCGGCGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTCGCC 80
QY 61 CGAGCGCCGGGCTTTGCCCGGCGGCGCTCACTAGCGAGCGAGCGCGAGAGGGAGTG 120
DB 81 CGAGCGCCGGGCTTTGCCCGGCGGCGCTCACTAGCGAGCGAGCGCGAGAGGGAGTG 140
QY 121 GCCAA 125
DB 141 GCCAA 145

RESULT 9
I16806
LOCUS
I16806
DEFINITION
Sequence 1 from patent US 5869305.
ACCESSION
AR034135
VERSION
AR034135.1 GI:5949740
KEYWORDS
Unknown.
SOURCE
Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 165)
AUTHORS
Samulski, R.Jude. and Xiao, X.
TITLE
Recombinant viral vector system
JOURNAL
Patent: US 5869305-A 1 09-FEB-1999;
FEATURES
Location/Qualifiers
1..165
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 125; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.7e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 TTGGCCACTCCTCTCTCGGCGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTCGCC 60
DB 21 TTGGCCACTCCTCTCTCGGCGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTCGCC 80
QY 61 CGAGCGCCGGGCTTTGCCCGGCGGCGCTCACTAGCGAGCGAGCGCGAGAGGGAGTG 120
DB 81 CGAGCGCCGGGCTTTGCCCGGCGGCGCTCACTAGCGAGCGAGCGCGAGAGGGAGTG 140
QY 121 GCCAA 125
DB 141 GCCAA 145

RESULT 9
I16806
LOCUS
I16806
DEFINITION
Sequence 1 from patent US 5869305.
ACCESSION
AR034135
VERSION
AR034135.1 GI:5949740
KEYWORDS
Unknown.
SOURCE
Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 165)
AUTHORS
Samulski, R.Jude. and Xiao, X.
TITLE
Recombinant viral vector system
JOURNAL
Patent: US 5869305-A 1 09-FEB-1999;
FEATURES
Location/Qualifiers
1..165
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 125; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.7e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 TTGGCCACTCCTCTCTCGGCGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTCGCC 60
DB 21 TTGGCCACTCCTCTCTCGGCGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTCGCC 80
QY 61 CGAGCGCCGGGCTTTGCCCGGCGGCGCTCACTAGCGAGCGAGCGCGAGAGGGAGTG 120
DB 81 CGAGCGCCGGGCTTTGCCCGGCGGCGCTCACTAGCGAGCGAGCGCGAGAGGGAGTG 140
QY 121 GCCAA 125
DB 141 GCCAA 145

RESULT 9
I16806
LOCUS
I16806
DEFINITION
Sequence 1 from patent US 5869305.
ACCESSION
AR034135
VERSION
AR034135.1 GI:5949740
KEYWORDS
Unknown.
SOURCE
Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 165)
AUTHORS
Samulski, R.Jude. and Xiao, X.
TITLE
Recombinant viral vector system
JOURNAL
Patent: US 5869305-A 1 09-FEB-1999;
FEATURES
Location/Qualifiers
1..165
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 125; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.7e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 TTGGCCACTCCTCTCTCGGCGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTCGCC 60
DB 21 TTGGCCACTCCTCTCTCGGCGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTCGCC 80
QY 61 CGAGCGCCGGGCTTTGCCCGGCGGCGCTCACTAGCGAGCGAGCGCGAGAGGGAGTG 120
DB 81 CGAGCGCCGGGCTTTGCCCGGCGGCGCTCACTAGCGAGCGAGCGCGAGAGGGAGTG 140
QY 121 GCCAA 125
DB 141 GCCAA 145

RESULT 9
I16806
LOCUS
I16806
DEFINITION
Sequence 1 from patent US 5869305.
ACCESSION
AR034135
VERSION
AR034135.1 GI:5949740
KEYWORDS
Unknown.
SOURCE
Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 165)
AUTHORS
Samulski, R.Jude. and Xiao, X.
TITLE
Recombinant viral vector system
JOURNAL
Patent: US 5869305-A 1 09-FEB-1999;
FEATURES
Location/Qualifiers
1..165
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 125; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.7e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 TTGGCCACTCCTCTCTCGGCGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTCGCC 60
DB 21 TTGGCCACTCCTCTCTCGGCGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTCGCC 80
QY 61 CGAGCGCCGGGCTTTGCCCGGCGGCGCTCACTAGCGAGCGAGCGCGAGAGGGAGTG 120
DB 81 CGAGCGCCGGGCTTTGCCCGGCGGCGCTCACTAGCGAGCGAGCGCGAGAGGGAGTG 140
QY 121 GCCAA 125
DB 141 GCCAA 145

RESULT 9
I16806
LOCUS
I16806
DEFINITION
Sequence 1 from patent US 5869305.
ACCESSION
AR034135
VERSION
AR034135.1 GI:5949740
KEYWORDS
Unknown.
SOURCE
Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 165)
AUTHORS
Samulski, R.Jude. and Xiao, X.
TITLE
Recombinant viral vector system
JOURNAL
Patent: US 5869305-A 1 09-FEB-1999;
FEATURES
Location/Qualifiers
1..165
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 125; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.7e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 TTGGCCACTCCTCTCTCGGCGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTCGCC 60
DB 21 TTGGCCACTCCTCTCTCGGCGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTCGCC 80
QY 61 CGAGCGCCGGGCTTTGCCCGGCGGCGCTCACTAGCGAGCGAGCGCGAGAGGGAGTG 120
DB 81 CGAGCGCCGGGCTTTGCCCGGCGGCGCTCACTAGCGAGCGAGCGCGAGAGGGAGTG 140
QY 121 GCCAA 125
DB 141 GCCAA 145

RESULT 9
I16806
LOCUS
I16806
DEFINITION
Sequence 1 from patent US 5869305.
ACCESSION
AR034135
VERSION
AR034135.1 GI:5949740
KEYWORDS
Unknown.
SOURCE
Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 165)
AUTHORS
Samulski, R.Jude. and Xiao, X.
TITLE
Recombinant viral vector system
JOURNAL
Patent: US 5869305-A 1 09-FEB-1999;
FEATURES
Location/Qualifiers
1..165
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 125; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.7e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 TTGGCCACTCCTCTCTCGGCGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTCGCC 60
DB 21 TTGGCCACTCCTCTCTCGGCGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTCGCC 80
QY 61 CGAGCGCCGGGCTTTGCCCGGCGGCGCTCACTAGCGAGCGAGCGCGAGAGGGAGTG 120
DB 81 CGAGCGCCGGGCTTTGCCCGGCGGCGCTCACTAGCGAGCGAGCGCGAGAGGGAGTG 140
QY 121 GCCAA 125
DB 141 GCCAA 145

RESULT 9
I16806
LOCUS
I16806
DEFINITION
Sequence 1 from patent US 5869305.
ACCESSION
AR034135
VERSION
AR034135.1 GI:5949740
KEYWORDS
Unknown.
SOURCE
Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 165)
AUTHORS
Samulski, R.Jude. and Xiao, X.
TITLE
Recombinant viral vector system
JOURNAL
Patent: US 5869305-A 1 09-FEB-1999;
FEATURES
Location/Qualifiers
1..165
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 125; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.7e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 TTGGCCACTCCTCTCTCGGCGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTCGCC 60
DB 21 TTGGCCACTCCTCTCTCGGCGCTCGCTCGCTCACTAGCGCGGCGACCAAGGTCGCC 80
QY 61 CGAGCGCCGGGCTTTGCCCGGCGGCGCTCACTAGCGAGCGAGCGCGAGAGGGAGTG 120
DB 81 CGAGCGCCGGGCTTTGCCCGGCGGCGCTCACTAGCGAGCGAGCGCGAGAGGGAGTG 140
QY 121 GCCAA 125
DB 141 GCCAA 145

RESULT 9
I16806
LOCUS
I1680
```

```
LOCUS AX106702 165 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 7 from Patent WO0125465.
ACCESSION AX106702
VERSION AX106702.1 GI:13922363
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS Engelhardt, J.F., Dongsheng, D. and Ziyang, Y.
TITLE Adeno-associated viruses and uses thereof
JOURNAL Patent: WO 0125465-A 7 12-APR-2001;
US University of Iowa Research Foundation (US); Engelhardt, John F.
(US); Dongsheng, Duan (US); Ziyang, Yan (US)
FEATURES
source
1. 165
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
/note="SEQ ID NO:1 of U.S. Patent No. 5,478,745"
ORIGIN
Query Match 100.0%; Score 125; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.7e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTAGCGCGCGCGACCAAGGTCGCC 60
DB 21 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTAGCGCGCGCGACCAAGGTCGCC 80
QY 61 CGAGCCCGGGCTTTGCCCGGCGCCCTCAGTGCAGCGAGCGCGCGAGAGGGAGTG 120
DB 81 CGAGCCCGGGCTTTGCCCGGCGCCCTCAGTGCAGCGAGCGCGCGAGAGGGAGTG 140
QY 121 GCCAA 125
DB 141 GCCAA 145
RESULT 12
LOCUS BD218219 165 bp DNA linear PAT 17-JUL-2003
DEFINITION Adeno-associated virus vectors and uses thereof.
ACCESSION BD218219
VERSION BD218219.1 GI:33027989
KEYWORDS JP 2002515257-A/7.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS Engelhardt, J.F., Duan, D. and Yang, T.
TITLE Adeno-associated virus vectors and uses thereof
JOURNAL Patent: JP 2002515257-A 7 28-MAY-2002;
UNIVERSITY OF IOWA RESEARCH FOUNDATION
COMMENT OS Unknown
FN JP 2002515257-A/7
PD 28-MAY-2002
PF 20-MAY-1999 JP 2000549752
PR 20-MAY-1998 US 60/086166, 25-MAR-1999 US 09/276625 PI
PC C12N15/00, A01K67/027, A61K31/711, A61K48/00, C07K14/47, C12N5/10,
PC C12N15/00,
PC C12N5/00
CC SEQ ID NO:1 of U.S. Patent No. 5,478,745
FH Key Location/Qualifiers
FT source 1. 165
FT /organism="Unknown".
FEATURES
source
1. 165
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 125; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.7e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTAGCGCGCGCGACCAAGGTCGCC 60
DB 21 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTAGCGCGCGCGACCAAGGTCGCC 80
QY 61 CGAGCCCGGGCTTTGCCCGGCGCCCTCAGTGCAGCGAGCGCGCGAGAGGGAGTG 120
DB 81 CGAGCCCGGGCTTTGCCCGGCGCCCTCAGTGCAGCGAGCGCGCGAGAGGGAGTG 140
QY 121 GCCAA 125
DB 141 GCCAA 145
RESULT 13
LOCUS A46400 192 bp DNA linear PAT 07-MAR-1997
DEFINITION Sequence 3 from Patent WO9523867.
ACCESSION A46400
VERSION A46400.1 GI:2300601
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 192)
AUTHORS Densfle, P., Latta, M., Perricaudet, M. and Vigne, E.
TITLE INTERACTIVE RECOMBINANT ADENOVIRUSES, PREPARATION THEREOF AND THERAPEUTICAL USES THEREOF
JOURNAL Patent: WO 9523867-A 3 08-SEP-1995;
RHONE POULENC RORER SA (FR)
COMMENT Other publication AU 1852695 950918
Other publication FR 2716893 950908.
FEATURES
source
1. 192
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 125; DB 6; Length 192;
Best Local Similarity 100.0%; Pred. No. 2.6e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTAGCGCGCGCGACCAAGGTCGCC 60
DB 68 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTAGCGCGCGCGACCAAGGTCGCC 127
QY 61 CGAGCCCGGGCTTTGCCCGGCGCCCTCAGTGCAGCGAGCGCGCGAGAGGGAGTG 120
DB 128 CGAGCCCGGGCTTTGCCCGGCGCCCTCAGTGCAGCGAGCGCGCGAGAGGGAGTG 187
QY 121 GCCAA 125
DB 188 GCCAA 192
RESULT 14
LOCUS AX418199/c 198 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 1 from Patent WO0192551.
ACCESSION AX418199
VERSION AX418199.1 GI:21523203
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Samulski, R.J. and Mccarty, D.M.
TITLE Duplexed parvovirus vectors
```


JOURNAL Patent: WO 0192551-A 1 06-DEC-2001;
UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL (US)
FEATURES Location/Qualifiers
source 1..198
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Inverted terminal repeat from the AAV-2 vector
plasmid pSub 201"

ORIGIN
Query Match 100.0%; Score 125; DB 6; Length 198;
Best Local Similarity 100.0%; Pred. No. 2.6e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGACCAAGGTGCGCC 60
Db |||||
125 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGACCAAGGTGCGCC 66
QY 61 CGACGCCCGGGGTTTGGCCCGCGGCGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
Db |||||
65 CGACGCCCGGGGTTTGGCCCGCGGCGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 6
QY 121 GCCAA 125
Db |||||
5 GCCAA 1

RESULT 15
AX703496
LOCUS AX703496 207 bp DNA linear PAT 03-APR-2003
DEFINITION Sequence 58 from Patent WO02066653.
ACCESSION AX703496
VERSION AX703496.1 GI:29538456
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Li, M. and Liu, Y. C.
TITLE Prokaryotic libraries and uses
JOURNAL Patent: WO 02066653-A 58 29-AUG-2002;
Xencor (US)
FEATURES Location/Qualifiers
source 1..207
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="synthetic enzyme attachment site sequence"

ORIGIN
Query Match 100.0%; Score 125; DB 6; Length 207;
Best Local Similarity 100.0%; Pred. No. 2.6e-18;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGACCAAGGTGCGCC 60
Db |||||
42 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGACCAAGGTGCGCC 101
QY 61 CGACGCCCGGGGTTTGGCCCGCGGCGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
Db |||||
102 CGACGCCCGGGGTTTGGCCCGCGGCGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 161
QY 121 GCCAA 125
Db |||||
162 GCCAA 166

Search completed: April 1, 2004, 16:04:10
Job time : 1574.61 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 1, 2004, 14:50:59 ; Search time 1603.24 Seconds
(without alignments)
2328.267 Million cell updates/sec

Title: US-10-620-039-1_COPY_1_125
Perfect score: 125
Sequence: 1 TTGGCCACTCCTCTCTGCG.....CGCAGAGAGGAGTGGCCAA 125

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estmu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_htc.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_hci.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: em_estfun.*
- 16: em_estom.*
- 17: em_gss_hum.*
- 18: em_gss_inv.*
- 19: em_gss_pln.*
- 20: em_gss_vrt.*
- 21: em_gss_fun.*
- 22: em_gss_nam.*
- 23: em_gss_mus.*
- 24: em_gss_pro.*
- 25: em_gss_rtd.*
- 26: em_gss_pbg.*
- 27: em_gss_vri.*
- 28: gb_gss1.*
- 29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	37.8	30.2	304	10	B2495968
C 2	37.4	29.9	1015	28	B2569259
C 3	37	29.6	1049	29	AG185368
4	36.6	29.3	644	13	BY752719

RESULT 1
BE495968/c
LOCUS
DEFINITION
CDNA clone WHE1260_H03_O06, mRNA sequence.
ACCESSION
BE495968
VERSION
EST.
KEYWORDS
SOURCE
ORGANISM
Scalae cereale (rye)
Scalae cereale
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Lillioopsida; Poales; Poaceae; Pooideae; Triticeae; Secale.
REFERENCE
1 (bases 1 to 304)
Fenton, R.D., Gustafson, J.P., Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J., Ross, K., Seaton, C.L. and Tong, J.C.
The structure and function of the expressed portion of the wheat Genomes - Another cdna library from rye
JOURNAL
Unpublished (2000)
COMMENT
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818

ALIGNMENTS

BE495968 304 bp mRNA linear EST 02-AUG-2000
WHE1260_H03_O06ZS Secale cereale anther cdna library
CDNA clone WHE1260_H03_O06, mRNA sequence.
BE495968
GI:9662561
Scalae cereale (rye)
Scalae cereale
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Lillioopsida; Poales; Poaceae; Pooideae; Triticeae; Secale.
1 (bases 1 to 304)
Fenton, R.D., Gustafson, J.P., Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J., Ross, K., Seaton, C.L. and Tong, J.C.
The structure and function of the expressed portion of the wheat Genomes - Another cdna library from rye
JOURNAL
Unpublished (2000)
COMMENT
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818

5	35.4	28.3	1103	13	BX403654
6	35.4	28.3	1182	12	EM807831
7	35	28.0	595	28	BH218838
8	35	28.0	896	13	BU152072
9	35	28.0	1103	13	BX403654
10	34.8	27.8	530	12	BM079047
11	34.6	27.7	514	13	BX338963
12	34.6	27.7	598	29	CG273564
13	34.6	27.7	654	28	B2532489
14	34.6	27.7	747	28	B2410270
15	34.6	27.7	827	29	CG285957
16	34.6	27.7	839	29	CG365468
17	34.6	27.7	840	29	CG441848
18	34.6	27.7	932	29	CNS00720
19	34.4	27.5	878	9	AL552821
20	34.2	27.4	969	9	AL564360
21	34	27.2	640	13	BY733594
22	34	27.2	933	29	AG060831
23	34	27.2	1121	29	CNS05GBJ
24	34	27.2	1123	29	AG080476
25	34	27.2	1201	13	BX405071
26	33.8	27.0	663	14	CD650788
27	33.8	27.0	1000	13	BX407619
28	33.8	27.0	1697	10	AW727558
29	33.6	26.9	430	29	CC778537
30	33.6	26.9	1057	29	CNS046BO
31	33.6	26.9	1057	29	CNS046BO
32	33.6	26.9	1471	28	BZ577163
33	33.4	26.7	433	14	CA697866
34	33.4	26.7	509	29	CG712785
35	33.4	26.7	514	13	BX338963
36	33.4	26.7	574	28	BH225085
37	33.4	26.7	578	28	BH218943
38	33.4	26.7	579	28	BH218776
39	33.4	26.7	594	28	BH218854
40	33.4	26.7	630	29	CNS012KV
41	33.4	26.7	808	28	BZ742499
42	33.4	26.7	1011	13	BQ961334
43	33.4	26.7	1049	12	BQ048852
44	33.4	26.7	1201	13	BX403541
45	33.2	26.6	366	13	BY303434

Email: candersn@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low
'quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.

FEATURES
source
1. .304
/organism="Secale cereale"
/mol_type="rRNA"
/cultivar="Blanco"
/db_xref="taxon:4550"
/clone="WHE1260 H03 O06"
/tissue_type="Anther"
/dev_stage="Adult plant before anthesis"
/lab_host="E. coli SOLR"
/clone_lib="Secale cereale anther cDNA library"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
site 1: EcoRI; Site 2: XhoI; Plants were grown in the
greenhouse. Anthers were harvested and pooled from early
meiosis to late meiosis. The tissue total RNA, and
poly(A) RNA were prepared (Butler, Ross and Gustafson) at
University of Missouri, Columbia. A cDNA library was
made, and the cDNA clones were in vivo excised to give
pBluescript phagemids in the TJ Close lab (Choi, Close,
Fenton) at the University of California, Riverside.
Plasmid DNA preparations and DNA sequencing were performed
in the OD Anderson lab (all other authors)."

ORIGIN
Query Match 30.2%; Score 37.8; DB 10; Length 304;
Best Local Similarity 57.9%; Pred. No. 36;
Matches 66; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
QY 2 TGGCCACTCCCTCTCGCGCTCGCTCCTACTGAGCGCGGCGGACCAAGGTGCGCC 61
DB 118 TTGCCACTCTCTTTTCACAGTCTACTACGACAGAGTAGGAGAGACACACTA 59
QY 62 GACGCCCGCGCTTCCCGCGGCGCTCAGTGAGCGAGCGCGGACGAGG 115
DB 58 GACACGAGATTTTTCGCGCCGACACGACGACGCGCGTGGCGCACAGC 5

RESULT 2
BZ569259/c
LOCUS
pac2-164_8319.x2r5p12 pac2-164 Pseudomonas aeruginosa genomic
clone pac2-164_8319, genomic survey sequence.
ACCESSION BZ569259
VERSION BZ569259.1 GI:27203589
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
REFERENCE Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
1 (bases 1 to 1015)
TITLE Whole-Genome-Sequence Variation among multiple isolates of
J. Bacteriol. (2002) In press
JOURNAL Contact: Chris K. Raymond
COMMENT Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES
source
1. .1015
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"

/clone="pac2-164_8319"
/clone_lib="pac2-164"
/note="clinical isolate 2-164 Whole genomic shotgun
library."

ORIGIN
Query Match 29.9%; Score 37.4; DB 28; Length 1015;
Best Local Similarity 56.7%; Pred. No. 61;
Matches 68; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 4 GCCACTCCCTCTCTCGCGCTCGCTCCTACTGAGCGCGGCGGACCAAGGTGCGCCGA 63
DB 952 GCCTCGCGGGTGGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 893
QY 64 CGCCCGGGCTTTCGCCCGCGCGCTCAGTGAGCGAGCGCGGACGAGAGGAGTGCGCC 123
DB 892 CG 833

RESULT 3
AG185368/c
LOCUS
Pan troglodytes DNA, clone: RP43-059G23.TJ, genomic survey
sequence.
DEFINITION
AG185368
ACCESSION AG185368.1 GI:16715048
VERSION AG185368
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE 1
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Toroki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library RPCI-43
TITLE Unpublished
JOURNAL 2 (bases 1 to 1049)
REFERENCE Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Toroki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22, Saitama-shi, Saitama, Japan; Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimps@sc.riken.go.jp, URL:http://hgp.riken.go.jp/;
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
end was generated during the R&D process and may have higher chance
of clone tracking errors.
PRIMERS
Sequencing: T1

LIBRARY
Vector : pBACe3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI.
Location/Qualifiers
1. .1049
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="RP43-059G23.TJ"
/sex="male"
/cell_type="lymphocytes"
/clone_lib="RPCI-43 Chimpanzee Male BAC Library"

ORIGIN
Query Match 29.8%; Score 37; DB 29; Length 1049;
Best Local Similarity 56.8%; Pred. No. 77;
Matches 67; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
QY 4 GCCACTCCCTCTCTCGCGCTCGCTCCTACTGAGCGCGGCGGACCAAGGTGCGCCGA 63
DB 527 CG 468

```

QY 64 CGCCCGGGCTTTGCCCGGGCGCTCAGTACGCGAGCGAGCGCGCAGAGGGAGTGG 121
    |||||
Db 467 CGCCCGGGCGGGGGGGGGGGTCTCCGCGTGGCGGGGGCGCGCTCCGCGGGGTGG 410
    |||||

RESULT 4
BY752719
LOCUS BY752719 RIKEN full-length enriched, adult inner ear Mus musculus
DEFINITION cDNA clone F930013F16 5', mRNA sequence.
ACCESSION BY752719
VERSION BY752719.1 GI:27183756
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
1 (bases 1 to 644)
AUTHORS Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
Schonbach,C., Gojobori,T., Schrinl,L.M., Kanapin,A., Matsuda,H.,
Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
Ciothia,C., Corbani,L.B., Cousins,S., Dalia,E., Dragani,T.A.,
Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
Kawaji,H., Kawasawa,Y., Kedzierzki,R.M., King,B.L., Konagaya,A.,
Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagasawa,T.,
Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
Verardo,R., Wagner,L., Wallestedt,C., Wang,Y., Watanabe,Y.,
Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
Rogers,J., Birney,E. and Hayashizaki,Y.
Contact: Yoshihide Hayashizaki
12466851
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp
URL: http://genome.gsc.riken.go.jp/

ADACHI,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,
Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,
Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,
Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,
Nomura,K., Numazaki,R., Ono,M., Ohsato,N., Saito,R., Sakazume,N.,
Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,
Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format

```

```

sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Kirk W. Beisel ( Boys Town National
Research Hospital 555 North 30th Street Omaha,NE 68131 USA ) whose
assistance we gratefully acknowledge.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

```

FEATURES

source

Location/Qualifiers

```

1..644
/organism="Mus musculus"
/mol_type="mRNA"
/strains="C57BL/6J"
/db_xref="taxon:10090"
/clone="F930013F16"
/tissue_type="inner ear"
/dev_stage="adult"
/clone_lib="RIKEN full-length enriched, adult inner ear"

```

ORIGIN

```

Query Match      29.3%; Score 36.6; DB 13; Length 644;
Best Local Similarity 56.1%; Pred. No. 84;
Matches 69; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 3  GGCCACTCCCTCTCTGCGGCTCGCTCAGTACGCGAGCGCGCGGAGGTGCGCCG 62
    |||||
Db 217 GGCCACTCCGCTATTTGGGCTTGTCTGTGGCGCTGGAGCGCGCGCGCGCGTGC 276

QY 63  AGCCCGGGCTTTGCCCGGGCGCTCAGTACGCGAGCGCGCGCGAGGGAGTGGC 122

Db 277 CCCCACACCTGTCTCCAGCGTCCCTGGAGCGCGCGCGCGCGACATGACGCTG 336

QY 123 CAA 125

Db 337 AAA 339

```

RESULT 5

BX403654

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

1..1103

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

BX403654

3-PRIME, mRNA sequence.

BX403654.1

EST.

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1103)

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/

Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CLOBA007ZH2FP1.

b

478 CGGCTGGCTTCTGTCCGGAAGCCGGGGTTCGCGGGGACCTCGAGATGAACGAG 424

```

RESULT 8
BUI52072
LOCUS       BUI52072      896 bp      mRNA      linear      EST 03-SEP-2002
DEFINITION  AGENCOURT_8749018 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6334507
5', mRNA sequence.
ACCESSION   BUI52072
VERSION     BUI52072.1  GI:22665604
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
REFERENCE   1
AUTHORS     Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE       Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL     NIH-MGC http://mgc.nci.nih.gov/.
COMMENT     National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
            Ph.D.
            cDNA Library Preparation: ResGen, Invitrogen Corp
            DNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LAM13795 row: 9 column: 20
            High quality sequence start: 3
            High quality sequence stop: 693.
            Location/Qualifiers
              1..896
                /organism="Mus musculus"
                /mol_type="mRNA"
                /db_xref="taxon:10090"
                /clone="IMAGE:6334507"
                /lab_host="DH10B (phage-resistant)"
                /clone_lib="NIH_MGC_130"
                /note="Organ: oocytes; Vector: pCMV-SPORT6.1; Site 1:
                EcorV; Site 2: NotI; Cloned unidirectionally. Primer:
                Oligo 48. Average insert size 1.95 kb. Constructed by
                ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ORIGIN
Query Match      28.0%; Score 35; DB 13; Length 896;
Best Local Similarity 55.3%; Pred. No. 2.2e+02;
Matches 68; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 3  GGCCACTCCCTCTGCGCGCTGCTGCTCACTGAGCGCGGCGACCAAGGTCGCGCG 62
      |||||
DB 557 GGCCACTTGGCTATTGGGCTTGTCTGTGCGCGTGGAGGGCCACGTCGCGGTC 616-
      |||||
QY 63  AGCGCCGGGCTTTCGCGGGCGGCTCACTGAGCGAGCGAGCGGCGAGAGGAGTGCG 122
      |||||
DB 617 CCGCGCACACTGTCCCGAGCGTCCGTCGGAGCGCGGCGGCGGCGACATGCGGCTG 676
      |||||
QY 123 CAA 125
      |||
DB 677 AAA 679

RESULT 9
BX403654/c
LOCUS       BX403654      1103 bp      mRNA      linear      EST 15-MAY-2003
DEFINITION  BX403654 Homo sapiens PLACENTA Homo sapiens cDNA clone CLOBA007ZH12
3'-PRIME, mRNA sequence.
ACCESSION   BX403654
VERSION     BX403654.1  GI:30762430
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1
AUTHORS     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE       Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL
COMMENT

```

```

REFERENCE   1 (bases 1 to 1103)
AUTHORS     Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CLOBA007ZH12FPI.
            Location/Qualifiers
              1..1103
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CLOBA007ZH12"
                /issue_type="PLACENTA"
                /clone_lib="Homo sapiens PLACENTA"
                /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
                with a NotI-oligo(dT) primer. Five prime end enriched,
                double-strand cDNA was digested with Not I and cloned into
                the Not I and EcoRV sites of the pCMVSPORT 6 vector.
                Library was not normalized."

ORIGIN
Query Match      28.0%; Score 35; DB 13; Length 1103;
Best Local Similarity 10.8%; Pred. No. 2.3e+02;
Matches 12; Conservative 69; Mismatches 30; Indels 0; Gaps 0;

QY 13  TCTTCGCGCTGCTGCTCACTGAGCGCGGCGACCAAGTCCCGAGCGCGCGGC 72
      :|||
DB 503 WSTCTCSSSSSTSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 444
      :|||
QY 73  TTTCGCGCGCGCTCACTGAGCGCGGCGGCGGAGGAGGAGTGCGGC 123
      :|||
DB 443 SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 393
      :|||

RESULT 10
BM079047/c
LOCUS       BM079047      530 bp      mRNA      linear      EST 14-NOV-2001
DEFINITION  MEST88-F06.T3 ISUM4-TN Zea mays cDNA clone MEST88-F06 3', mRNA
sequence.
ACCESSION   BM079047
VERSION     BM079047.1  GI:16925979
KEYWORDS    EST.
SOURCE      Zea mays
ORGANISM    Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 530)
Expressed Sequence Tags from B73 Maize Seedlings and Silks
Unpublished (2001)
Contact: Patrick S. Schnable
Schnable Laboratory
Iowa State University
G405 Agronomy, Iowa State University, Ames, IA 50011-1010, USA
Tel: 515-294-0975
Fax: 515-294-3299
Email: schnable@iastate.edu
Individual basecall and confidence value were assigned using the
Phred software.
(<http://depts.washington.edu/ventures/collabtr/direct/index.htm#b
rt>). Overall sequence quality assessment and vector trimming were
conducted using the Lucy software (<http://www.tigr.org/softlab/>).
Lucy parameters were set to ensure an overall trimmed quality of
97.5% or better without any vector fragments in the chosen
high-quality region of each sequence. Low-quality bases between the
poly-T and the high-quality region were replaced with N's to serve

```



```

Db 169 GTGAGCTTGGTGAAGCACTTGAGCGGGCAACCTCCGCCAGTATGGAGCAGCAGGATGGC 228
QY 78 CGCGCGCGCTCAGTGAGCGAGCGAGCGCGGAGAG 114
Db 229 CTGGCGCGCACCATTTGAGAGCTCGAGCTCCTTGAGAG 265

RESULT 13
BZ532489
LOCUS OGAEY44TC ZM2.0.7.1.5_KB Zea mays genomic clone ZMMBMA0046G15,
DEFINITION genomic survey sequence.
ACCESSION BZ532489
VERSION BZ532489.1 GI:27077191
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 654)
REFERENCE
AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
CONTACT: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..654
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0046G15"
/clone_lib="ZM2.0.7.1.5_KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 27.7%; Score 34.6; DB 28; Length 654;
Best Local Similarity 59.8%; Pred No. 2.5e+02;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 18 GCGCGCTCGCTCAGTGAAGCGGGCGACCAAGTGCAGCCGCGGCTTGC 77
Db 513 GTGAGCTTGGTGAAGCACTTGAGCGGGCAACCTCCGCCAGTATGGAGCAGCAGGATGGC 572
QY 78 CGCGCGCGCTCAGTGAGCGAGCGAGCGCGGAGAG 114
Db 573 CTGGCGCGCACCATTTGAGAGCTCGAGCTCCTTGAGAG 609

RESULT 14
BZ410270/c
LOCUS OGAB801TC ZM0.7.1.5_KB Zea mays genomic clone ZMMBMA0020023,
DEFINITION genomic survey sequence.
ACCESSION BZ410270
VERSION BZ410270.1 GI:26042544
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 747)
REFERENCE
AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
CONTACT: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..747
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0020023"
/clone_lib="ZM0.7.1.5_KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 27.7%; Score 34.6; DB 28; Length 747;
Best Local Similarity 59.8%; Pred No. 2.6e+02;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 18 GCGCGCTCGCTCAGTGAAGCGGGCGACCAAGTGCAGCCGCGGCTTGC 77
Db 593 GTGAGCTTGGTGAAGCACTTGAGCGGGCAACCTCCGCCAGTATGGAGCAGCAGGATGGC 534
QY 78 CGCGCGCGCTCAGTGAGCGAGCGAGCGCGGAGAG 114
Db 533 CTGGCGCGCACCATTTGAGAGCTCGAGCTCCTTGAGAG 497

RESULT 15
CG285957/c
LOCUS OGWK933TH ZM0.7.1.5_KB Zea mays genomic clone ZMMBMA0607P17,
DEFINITION genomic survey sequence.
ACCESSION CG285957
VERSION CG285957.1 GI:34200171
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 827)
REFERENCE
AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OGWK93TV
CONTACT: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..827
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBMA0607P17"

```


/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCK-; Site_1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN

Query Match	27.7%	Score 34.6;	DB 29;	Length 827;
Best Local Similarity	59.8%	Pred. No. 2.7e+02;		
Matches	58;	Conservative	0;	Mismatches 39;
				Indels 0;
				Gaps 0;

QY	18	GCGCGCTCGCTCGCTCACTGAGCGCGGGCGACCAAGGTCGCCCGCGCGGCTTTGC	77
Db	351	GTGAGCTTGGTGAGGCACTTGAGCGCGGCAACCTCCGCCAGTATGGAGCAGGATGGC	292

QY	78	CCGGCGGCTCAGTGAGCGAGCGCGCGAGAGAG	114
Db	291	CTGGCGGCAACCAATTGAGACTCGACTCTTGAGAG	255

Search completed: April 1, 2004, 17:02:17
Job time : 1606.24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 1, 2004, 14:50:59 ; Search time 1859.76 Seconds
(without alignments)
2328.267 Million cell updates/sec

Title: US-10-620-039-1
Perfect score: 145
Sequence: 1 TTGCCACTCCTCTCTGG.....CTCCATCACTAGGGTTCCT 145

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthu:*

3: em_estin:*

4: em_estm:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hic:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rtd:*

26: em_gss_pig:*

27: em_gss_vrl:*

28: gb_ges1:*

29: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	38.6	26.6	974	14	CF784470
C 2	37.8	26.1	304	10	BE495968
C 3	37.4	25.8	1015	28	BZ569259
C 4	37	25.5	1049	29	AG185368

5	36.6	25.2	644	13	BY752719
6	36.4	25.1	969	9	AL564360
C 7	35.8	24.7	1121	29	CNS05GBJ
C 8	35.6	24.6	1103	13	EX403654
9	35.6	24.6	1201	13	EX405071
10	35.4	24.4	598	28	CG273564
11	35.4	24.4	554	28	BZ534489
C 12	35.4	24.4	747	28	BZ410270
C 13	35.4	24.4	827	29	CG285957
14	35.4	24.4	839	29	CG365468
15	35.4	24.4	840	29	CG441848
16	35.4	24.4	1103	13	EX403654
17	35.4	24.4	1182	12	BM807831
18	35.2	24.3	1349	11	AK017628
C 19	35	24.1	595	28	BZ18838
20	35	24.1	896	13	BUL52072
21	35	24.1	1303	28	BZ570288
C 22	34.8	24.0	530	12	BM079047
C 23	34.8	24.0	953	13	BQ225579
C 24	34.6	23.9	514	13	EX338963
C 25	34.6	23.9	813	28	AZ193936
C 26	34.6	23.9	932	29	CNS0072Q
C 27	34.4	23.7	878	9	AL552621
C 28	34.2	23.6	1049	12	BQ48852
C 29	34.2	23.6	1201	13	EX360624
C 30	34	23.4	438	12	BM500627
31	34	23.4	640	13	BY733594
32	34	23.4	933	29	AG060831
33	34	23.4	1000	13	EX407619
34	34	23.4	1101	29	CNS017V2
35	34	23.4	1115	12	BM924606
36	34	23.4	1123	29	AG080476
37	33.8	23.3	663	14	CD650788
C 38	33.8	23.3	927	28	CG390088
39	33.8	23.3	1057	29	CNS046BO
40	33.8	23.3	1897	10	AW272558
41	33.6	23.2	430	29	CC778537
C 42	33.6	23.2	650	12	B1549023
C 43	33.6	23.2	874	13	EX389389
C 44	33.6	23.2	891	12	B1551983
C 45	33.6	23.2	1057	29	CNS046BO

ALIGNMENTS

RESULT 1
CF784470/c
LOCUS
DEFINITION
AGENCY: 15918074 X-St10-30 Silurana tropicalis cDNA clone
IMAGE: 7027270 3', mRNA sequence.
CF784470
CF784470.1 GI:37748315
EST.
Silurana tropicalis (western clawed frog)
Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 974)
NIH-MGC <http://imgc.ncbi.nih.gov/>.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
COMMENT
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Dr. Wei Wu / Prof. Christof NIEHR
CDNA Library Preparation: Wei Wu, DKPZ
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

CF784470 974 bp mRNA linear EST 20-OCT-2003
AGENCY: 15918074 X-St10-30 Silurana tropicalis cDNA clone
IMAGE: 7027270 3', mRNA sequence.

CF784470.1 GI:37748315

EST.

Silurana tropicalis (western clawed frog)

Silurana tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

Xenopodinae; Silurana.

1 (bases 1 to 974)

NIH-MGC <http://imgc.ncbi.nih.gov/>.

AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE

Unpublished (1999)

JOURNAL

COMMENT

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Dr. Wei Wu / Prof. Christof NIEHR

CDNA Library Preparation: Wei Wu, DKPZ

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM14761 row: d column: 20

High quality sequence start: 36

High quality sequence stop: 157

FEATURES

Location/Qualifiers
1..974
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:7027270"
/tissue_type="whole embryo, pool of stages 10, 20 and 30"
/clone_lib="Xtst10-30"
/note="Vector: PRW2; Site 1: BamHI; Site 2: XhoI; 10 ug of polyA+ RNA was isolated from a mixture of embryos at stage 10, 20 and 30 and primed by oligo-dT primer: 5'-GAGAGAGAGAGATCC(T)16VN-3' (where V=G,A,C). 5-methyl-dCTP was used instead of dCTP in the first-strand synthesis in order to get hemimethylated cDNA. After full-length enrichment, oligo-dG tailing and normalization against itself, second-strand synthesis was carried out by priming with 5'-GAGAGAGACTCGAGTTAATTAAT(C)13-3'. dsDNA was digested with XhoI/BamHI and directionally cloned into the pRKT2 vector. Average insert size is 1.5 kb. Library constructed using the Carninci protocol (Genome Research 2000) by Drs. W. Wu and C. Niehrs (DFPZ, Heidelberg, Germany)."

ORIGIN

Query Match 26.6%; Score 38.6; DB 14; Length 974;
Best Local Similarity 91.1%; Pred. No. 35;
Matches 41; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 101 GAGCGCGAGAGGGGCGCACTCCATCCTAGGGGTTCT 145

Db 159 GTGCATTCAGAGGGGAGTGGCACTCCATCCTAGGGGTTCT 115

RESULT 2

BE495968/c
LOCUS
DEFINITION
WHE1260_H03_006ZS Secale cereale anther cDNA library Secale cereale
cDNA clone WHE1260_H03_006, mRNA sequence.

ACCESSION
BE495968

VERSION
EST.

KEYWORDS
1 GI:9662561

SOURCE
Secale cereale (rye)

ORGANISM
Secale cereale

REFERENCE
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Authors
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Keywords
Poaceae; Triticeae; Secale.

1 (bases 1 to 304)

Anderson, O.D., Butler, E., Chao, S., Choi, D.W., Close, T.J.,

Fenton, R.D., Gustafson, J.P., Han, P.S., Hsia, C.C., Kang, Y.,

Lazo, G.R., Miller, R., Rausch, C.J., Ross, K., Seaton, C.L. and

Tong, J.C.

The structure and function of the expressed portion of the wheat

Genomes - Anther cDNA library from rye

Unpublished (2000)

Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific

West Area, Western Regional Research Center

800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105595773

Fax: 5105595818

Email: andersowp.usda.gov

Sequence have been trimmed to remove vector sequence and low

quality sequence with phred score less than 20

Seq primer: Stratagene SK primer.

Location/Qualifiers

1..304

/organism="Secale cereale"

/mol_type="mRNA"

FEATURES

source

Query Match 25.8%; Score 37.4; DB 28; Length 1015;

Best Local Similarity 56.7%; Pred. No. 68;

/cultivar="Blanco"
/db_xref="taxon:4550"
/clone="WHE1260_H03_006"
/tissue_type="Anther"
/dev_stage="Adult plant before anthesis"
/lab_host="E. coli SOLR"
/clone_lib="Secale cereale anther cDNA library"
/notes="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; Plants were grown in the
greenhouse. Anthers were harvested and pooled from early
meiosis to late meiosis. The tissue, total RNA, and
poly(A) RNA were prepared (Butler, Ross and Gustafson) at
University of Missouri, Columbia. A cDNA library was
made, and the cDNA clones were in vivo excised to give
pBluescript phagemids in the T3 Clone lab (Choi, Close,
Fenton) at the University of California, Riverside.
Plasmid DNA preparations and DNA sequencing were performed
in the OD Anderson lab (all other authors)."

ORIGIN

Query Match 26.1%; Score 37.8; DB 10; Length 304;
Best Local Similarity 57.9%; Pred. No. 38;
Matches 66; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 2 TGGCCACTCCCTCTCTCGCGCTCGCTCAGTCTAGCGCGCGACCAAGGTGCCCC 61

Db 118 TTGCCACTCTCTTCTCAGAGCTCACTCAGGACAGAGGTAGGAGAGACACACTA 59

QY 62 GAGCCCGGCTTTGCCGCGCGCTCAGTCTAGCGCGCGCGAGCGCGCAGAGAGG 115

Db 58 GACACGAGGATTTTTCGCGNCCGCGCACAGACCGCGGTGGCGGCACAGACG 5

RESULT 3

BZ569259/c

LOCUS

DEFINITION

pacS2-164_8319.x2r5p12 pacS2-164 Pseudomonas aeruginosa genomic

clone pacS2-164_8319, genomic survey sequence.

ACCESSION

BZ569259

KEYWORDS

BZ569259.1 GI:27203589

SOURCE

GSS.

Pseudomonas aeruginosa

ORGANISM

Pseudomonas aeruginosa

REFERENCE

1 (bases 1 to 1015)

Authors

Burns, J.L., Kaul, R. and Olsen, M.V.

Whole-Genome-Sequence variation among multiple isolates of

Pseudomonas aeruginosa library

J. Bacteriol. (2002) In press

CONTACT: Chris K. Raymond

Genome Center

University of Washington

Box 352145, Seattle, WA 98105-2145, USA

Tel: 2062216954

Fax: 2066857244

Email: craymond@u.washington.edu

Class: Shotgun.

Location/Qualifiers

1..1015

/organism="Pseudomonas aeruginosa"

/mol_type="genomic DNA"

/strain="2-164"

/db_xref="taxon:287"

/clone_lib="pacS2-164_8319"

/notes="clinical isolate 2-164 Whole genomic shotgun

library."

ORIGIN

Query Match 25.8%; Score 37.4; DB 28; Length 1015;

Best Local Similarity 56.7%; Pred. No. 68;

```

Matches 68; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
Qy 4 GCACCTCCTCTCTGCGGCTCGCTCCTCAGCGCCGCGGCGACCAAGGTCGCCCGA 63
Db 952 GCTTCGCGCGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 893
Qy 64 CGCCCGGCGCTTGGCCCGGCGGCTCAGTGAAGCGAGCGCGCGCAGAGAGGAGTGCC 123
Db 892 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCC 833

RESULT 4
AG185368/c
LOCUS
DEFINITION
  Pan troglodytes DNA, clone: RP43-059G23.TJ, genomic survey
  sequence.
ACCESSION
  AG185368
VERSION
  AG185368.1 GI:16715048
KEYWORDS
  GSS.
ORGANISM
  Pan troglodytes (chimpanzee)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE
  1 Fujiyama A., Hattori M., Toyoda A., Taylor T.D., Yada T.,
  Totoki Y., Watanabe H. and Sakaki Y.
  BAC end sequences of Library RPCI-43
  Unpublished
  2 (bases 1 to 1049)
  Fujiyama A., Hattori M., Toyoda A., Taylor T.D., Yada T.,
  Totoki Y., Watanabe H. and Sakaki Y.
  Direct Submission
  Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
  and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
  1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
  (E-mail: chimpanzee@gs.riken.go.jp, URL: http://hgpc.gsc.riken.go.jp/,
  Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT
  Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
  end was generated during the R&D process and may have higher chance
  of clone tracking errors.
PRIMERS
  Sequencing: TJ
LIBRARY
  Vector : pBACe3.6
  R.Site 1 : EcoRI
  R.Site 2 : EcoRI
  Location/Qualifiers
    1..1049
    /organism="Pan troglodytes"
    /mol_type="genomic DNA"
    /db_xref="taxon:9598"
    /clone="RP43-059G23.TJ"
    /sex="male"
    /cell_type="lymphocytes"
    /clone_lib="RPCI-43 Chimpanzee Male BAC Library"
ORIGIN
  Query Match 25.5%; Score 37; DB 29; Length 1049;
  Best Local Similarity 56.8%; Pred. No. 86;
  Matches 67; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
  Qy 4 GCACCTCCTCTCTGCGGCTCGCTCCTCAGCGCCGCGGCGACCAAGGTCGCCCGA 63
  Db 527 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 468
  Qy 64 CGCCCGGCGCTTGGCCCGGCGGCTCAGTGAAGCGAGCGCGCGCAGAGAGGAGTGCG 121
  Db 467 CGCCCGCGCGCGCGCGCGCGCTCTCGCGCTGCGCGGCGCGCGCGCGCGGTGG 410

RESULT 5
BY752719
LOCUS

```

```

DEFINITION
  BY752719 RIKEN full-length enriched, adult inner ear Mus musculus
  cDNA clone F930013F16 5', mRNA sequence.
ACCESSION
  BY752719
VERSION
  BY752719.1 GI:27193756
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Mus musculus
  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
  1 (bases 1 to 644)
  Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
  Niikado I., Oeato N., Saito R., Suzuki H., Yamataka I.,
  Kiyosawa H., Yagi K., Tomaru Y., Hasegawa Y., Nogami A.,
  Schonbach C., Gofjoberi T., Baldarelli R., Hill D.P., Bult C.,
  Hume D.A., Quackenbush J., Schriml L.M., Kanapin A., Matsuda H.,
  Batalov S., Beisel K.W., Blake J.A., Bradt D., Brusic V.,
  Chochia C., Corbani L.E., Cousins S., Della E., Dragani T.A.,
  Fletcher C.F., Forrest A., Frazer K.S., Gaasterland T.,
  Gariboldi M., Giesi C., Godzik A., Gough J., Grimmond S.,
  Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D., Kanai A.,
  Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L., Konagaya A.,
  Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A., Maglott D.R.,
  Maltais L., Marchionni L., McKenzie L., Miki H., Nagashima T.,
  Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
  Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
  Ravasi F., Reed J.C., Reid D.J., Reid J., Ring B.Z., Ringwald M.,
  Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
  Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
  Verdaro R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y.,
  Wells C., Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I.,
  Yang L., Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P.,
  Hayatsu N., Hirozane-Kishikawa T., Konno H., Nakamura M.,
  Sakazume N., Sato K., Shiraki T., Waki K., Kawai J., Aikawa K.,
  Aikawa T., Fukuda S., Hara A., Hashizume W., Imotani K., Ishii Y.,
  Itoh M., Kagawa I., Miyazaki A., Sakai K., Sasaki D., Shibata K.,
  Shinagawa A., Yasunishi A., Yoshino M., Waterston R., Lander E.S.,
  Rogers J., Birney E. and Hayashizaki Y.
  Analysis of the mouse transcriptome based on functional annotation
  of 50,770 full-length cDNAs
  Nature 420, 563-573 (2002)
JOURNAL
  MEDLINE
  PUBMED
  12466851
COMMENT
  Contact: Yoshihide Hayashizaki
  Laboratory for Genome Exploration Research Group, RIKEN Genomic
  Sciences Center (GSC), Yokohama Institute
  The Institute of Physical and Chemical Research (RIKEN)
  1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
  Tel: 81-45-503-9222
  Fax: 81-45-503-9216
  Email: genome-res@gs.riken.go.jp,
  URL: http://genome.gsc.riken.go.jp/
  Adachi J., Aizawa K., Akimura T., Arakawa T., Carninci P.,
  Fukuda S., Hashizume W., Hayashida K., Hirozane T., Hori F.,
  Imotani K., Ishii Y., Itoh M., Kagawa I., Kawai J., Kojima Y.,
  Kondo S., Konno H., Koya S., Miyazaki A., Murata M., Nakamura M.,
  Nomura K., Numazaki R., Ohno M., Ohsato N., Saito R., Sakazume N.,
  Sano H., Sasaki P., Sato K., Shibata K., Shiraki T., Tagami M.,
  Takeda Y., Waki K., Watahiki A., Muramatsu M. and Hayashizaki Y.
  Direct Submission
  Computational Analysis of Full-length Mouse cDNAs Compared with
  Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
  Normalization and subtraction of cap-trapper-selected cDNAs to
  prepare full-length cDNA libraries for rapid discovery of new
  genes. Genome Res. 10 (10), 1617-1630 (2000)
  RIKEN integrated sequence analysis (RISA) system--384-format
  sequencing pipeline with 384 multicapillary sequencer. Genome Res.
  10 (11), 1757-1771 (2000)
  Computer-based methods for the mouse full-length cDNA
  encyclopedia: real-time sequence clustering for construction of a
  nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
  cDNA library was prepared and sequenced in Mouse Genome
  Encyclopedia Project of Genome Exploration Research Group in Riken
  Genomic Sciences Center and Genome Science Laboratory in RIKEN.

```



```

/clone="003C03"
/clone.lib="B"
/notes="Genoscope sequence ID : COAB003AB02C1-end : T7"

ORIGIN

Query Match      24.7%; Score 35.8; DB 29; Length 1121;
Best Local Similarity 47.1%; Pred. No. 1.7e+02;
Matches 64; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 4 GCCACTCCTCTCTGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCCGCA 63
DB 664 GCGCGCNCNNGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 605

QY 64 GCGCGCGGCTTTGCGCGGCGCGCTCACTAGTACGAGCGCGCGCGAGAGGAGTGCC 123
DB 604 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 545

QY 124 AACTCCATCACTAGG 139
DB 544 GGCNCNCGCGCGCG 529

RESULT 8
BX403654/c
LOCUS BX403654 Homo sapiens PLACENTA mRNA linear EST 15-MAY-2003
DEFINITION 3-PRIME, mRNA sequence.
ACCESSION BX403654
VERSION BX403654.1 GI:30762430
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (Bases 1 to 1103)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CLOBA007ZH12FP1.
Location/Qualifiers
1..1103
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBA007ZH12"
/tissue_type="PLACENTA"
/clone.lib="Homo sapiens PLACENTA"
/notes="Vector: PCWVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match      24.6%; Score 35.6; DB 13; Length 1103;
Best Local Similarity 12.3%; Pred. No. 1.9e+02;
Matches 15; Conservative 71; Mismatches 36; Indels 0; Gaps 0;

QY 13 TCTCTGCGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCCGCGCGCGCGCG 72
DB 503 WSTCTGSSSSSSSTSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 444

QY 73 TTGCGCGGCGGCTCACTAGTACGAGCGCGCGAGAGGAGTGCCCACTCCATC 132
DB 443 SSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 384

```

```

QY 133 AC 134
DB 383 CC 382

RESULT 9
BX405071
LOCUS BX405071 Homo sapiens B CELLS (RAMOS CELL LINE) Homo sapiens cDNA
DEFINITION clone CS0D006YG06 3-PRIME, mRNA sequence.
ACCESSION BX405071
VERSION BX405071.1 GI:30648111
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1201)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10245.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0AG006BD03NP1&cluster=10245.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0AG006BD03NP1.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D006YG06"
/tissue_type="B CELLS (RAMOS CELL LINE)"
/clone.lib="RAMOS CELL LINE"
/notes="Vector: PCWVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match      24.6%; Score 35.6; DB 13; Length 1201;
Best Local Similarity 43.8%; Pred. No. 1.9e+02;
Matches 56; Conservative 19; Mismatches 53; Indels 0; Gaps 0;

QY 3 GGCACCTCCTCTCTGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCCG 62
DB 741 GGCACMGACACCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 800

QY 63 AGCGCGGCGCTTTGCGCGGCGCGCTCACTAGTACGAGCGCGCGAGAGAGTGCG 122
DB 801 SCGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 860

QY 123 CAACCTCCA 130
DB 861 TGGCGCCA 868

RESULT 10
CG273564
LOCUS CG273564 Zm_0.7_1.5_KB Zea mays genomic clone ZM5BMA075D08,
DEFINITION genomic survey sequence.
ACCESSION CG273564
VERSION CG273564.1 GI:34185705

```

GSS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE Zea mays
Zeakaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

AUTHORS Resnick,A., Fraser,C.M., Budinan,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Contact: Cathy Whitelaw

TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Contact: Cathy Whitelaw

FEATURES Location/Qualifiers
source 1..598
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBMA075D008"
 /clone_lib="ZM 0.7-1.5_KB"
 /note="Vector: pBCSK; Site 1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN
Query Match 24.4%; Score 35.4; DB 29; Length 598;
Best Local Similarity 57.8%; Pred No. 1.7e+02;
Matches 83; Conservative 0; Mismatches 46; Indels 0; Gaps 0

QY 18 GC CGCGTTCGCTCGCTACATGAGCGGGCGGCACCAGAGTGCCGCCGCGCGCTTGGC 77
Db 169 GTGAGCTTGCTGAGGCACTTGAGCGGGCACTCCGCCAGTATGGACGACGAGGTGC 228

QY 78 CCGGGGGCGCTCAGTGAGCGAGCGCGCAGAGGGAGTGCCCAAC 126
Db 229 CTGGGGCGCACCATTTGAGGCTCGAGCTCTTGAGACCTGCATCCGAC 277

RESULT 11
BZ532489 654 bp DNA linear GSS 16-DEC-2000
LOCUS OGAEY447c ZW2 0.7-1.5 kb Zea mays genomic clone ZMMBMA046G15,
DEFINITION genomic survey sequence.

ACCESSION BZ532489
VERSION BZ532489.1 GI:27077191
KEYWORDS GSS.

SOURCE Zea mays
ORGANISM Zea mays

REFERENCE Zeakaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budinan,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Contact: Cathy Whitelaw

TITLE Consortium for Maize Genomics
JOURNAL Unpublished (2002)
COMMENT Contact: Cathy Whitelaw

FEATURES Location/Qualifiers
source 1..598
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBMA075D008"
 /clone_lib="ZM 0.7-1.5_KB"
 /note="Vector: pBCSK; Site 1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

Db 533 CTGGCGCGCACCATGTAGAGCTCGAGTCTCTTGAGAGCGCTGCATCCGAC 485

CG285957 827 bp DNA linear GSS 25-AUG-2003
 CGWK933TH ZM 0.7 1.5 KB Zea mays genomic clone ZMMBma0607P17,
 genomic survey sequence.
 CG285957
 CG285957.1 GI:34200171
 GSS.
 Zea mays
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 827)
 Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
 Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
 Consortium for Maize Genomics
 Unpublished (2002)
 Other GSSs: OGWK933TV
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TR
 Class: sheared ends.

FEATURES
 source 1..827
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone="ZMMBma0607P17"
 /clone_lib="ZM 0.7 1.5 KB"
 /note="Vector: pBOSK; Site 1: HincII; 0.7-1.5 kb
 methylation filtered genomic DNA library"

ORIGIN
 Query Match 24.4%; Score 35.4; DB 29; Length 827;
 Best Local Similarity 57.8%; Pred. No. 1.9e-02;
 Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0

Qy 18 GCGCGCTCGCTCGCTCACTGAGCGCGGACCAAGCTCCCGACGCCGGGCTTGC 77
 Db 351 GTGAGCTTGTTGAGGCATTTAGCCGGGCAACCTCCGCAGTATGACACGAGATGCC 292

Qy 78 CCGGCGCGCCCTCAGTAGCGAGCGCGGACGAGAGGAGGTGGCCAC 126
 Db 291 CTGGCGCGCACCATGTAGAGCTCGAGTCTCTTGAGAGCTGCATCCGAC 243

RESULT 14
 CG365468 839 bp DNA linear GSS 26-AUG-2003
 LOCUS CG2BP05TH ZM 0.7 1.5 KB Zea mays genomic clone ZMMBma0751B10,
 genomic survey sequence.
 CG365468
 CG365468.1 GI:34282735
 GSS.
 Zea mays
 Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 839)
 Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
 Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
 Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.

REFERENCE
 AUTHORS

methylation filtered genomic DNA library"

ORIGIN

	Query Match	24.4%	Score 35.4	DB 29	Length 840
	Best Local Similarity	57.8%	Pred. No. 1.9e+02		
	Matches 63	Conservative 0	Mismatches 46	Indels 0	Gaps 0
Qy	18	GCGCGCTCGCTGGCTCACTGAGCGCGGGCGACCAAGTGC	CGCCGAGCGCCGGCGCTTGC	77	
Db	285	GTGAGCTTGTGTAGGCACTTGAGCGCGGCACTTCCGCCAGTATG	GAGACAGAGATGCG	344	
Qy	78	CCGGCGCGGCTCTACTGAGCGAGCGCGCAGAGGGGAGTGCCCAAC	126		
Db	345	CTGGGCGGCACCATTTGAGAGCTCGAGTCTTTGAGAGCGCTGCATCCGCAC	393		

Search completed: April 1, 2004, 17:02:14
Job time : 1875.76 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 1, 2004, 16:04:14 ; Search time 218.574 Seconds
(without alignments)
2480.895 Million cell updates/sec

Title: US-10-620-039-1

Perfect score:

Sequence: 1 TTGGCCACTCCCTCTCTGCG.....CTCCATCACTAGGGGTTCTT 145

Scoring table: IDENTITY_NUC

Searched: 2465228 seqs, 1869859620

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% .

Maximum Match 100%

[illegible]

```

database ;
  furnished_applications_mn:
  1: /cgn2_6/ptodata/1/pubpna

```

- 2: /cgn2_6/ptodata/1/pubpna/pct_NEW PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10_NEW PUB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US06_NEW PUB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query #			DB	ID	Description
		Match	Length				
1	1.45	100.0	145	9	US-09-782-378A-6	Sequence 6, Appli	
2	1.45	100.0	146	13	US-10-135-884-8	Sequence 8, Appli	
3	1.15	100.0	165	9	US-09-732-378A-8	Sequence 8, Appli	
4	1.45	100.0	165	13	US-10-054-665-7	Sequence 7, Appli	
C 5	1.45	100.0	165	14	US-10-259-568-13	Sequence 13, Appli	
C 6	1.45	100.0	175	12	US-10-276-356-1	Sequence 1, Appli	
7	1.45	100.0	207	14	US-10-023-008-58	Sequence 58, Appli	
8	1.45	100.0	955	10	US-09-845-416-26	Sequence 26, Appli	
C 9	1.45	100.0	955	10	US-09-845-416-33	Sequence 33, Appli	
10	1.45	100.0	987	10	US-09-845-416-33	Sequence 33, Appli	
C 11	1.45	100.0	987	10	US-09-845-416-33	Sequence 33, Appli	
12	1.45	100.0	4414	10	US-09-845-416-32	Sequence 32, Appli	
C 13	1.45	100.0	4414	10	US-09-845-416-32	Sequence 32, Appli	
C 14	1.45	100.0	4476	10	US-09-845-416-31	Sequence 31, Appli	
C 15	1.45	100.0	4476	10	US-09-845-416-31	Sequence 31, Appli	

16	145	100.0	4498	10	US-09-845-416-30	Sequence 30, Appl
17	145	100.0	4498	10	US-09-845-416-30	Sequence 30, Appl
18	145	100.0	4575	9	US-09-782-378A-1	Sequence 1, Appl
19	145	100.0	4575	9	US-09-782-378A-2	Sequence 2, Appl
20	145	100.0	4675	14	US-10-240-138-1	Sequence 1, Appl
21	145	100.0	4675	14	US-10-291-583-7	Sequence 7, Appl
22	145	100.0	4579	9	US-09-804-898-1	Sequence 1, Appl
23	145	100.0	4579	9	US-09-845-681-10	Sequence 10, Appl
24	145	100.0	4579	13	US-10-038-972A-12	Sequence 12, Appl
25	145	100.0	4579	14	US-10-135-813-6	Sequence 6, Appl
26	145	100.0	4680	13	US-10-077-284-1	Sequence 1, Appl
27	145	100.0	4680	13	US-10-163-886-1	Sequence 1, Appl
28	145	100.0	4680	14	US-10-263-137-1	Sequence 1, Appl
29	145	100.0	4680	14	US-10-375-777-1	Sequence 1, Appl
30	145	100.0	4681	12	US-10-695-261-18	Sequence 18, Appl
31	145	100.0	4681	12	US-10-695-282-18	Sequence 18, Appl
32	145	100.0	4681	12	US-10-695-900-18	Sequence 18, Appl
33	145	100.0	4683	12	US-10-695-281-19	Sequence 19, Appl
34	145	100.0	4683	12	US-10-695-282-19	Sequence 19, Appl
35	145	100.0	4683	12	US-10-695-900-19	Sequence 19, Appl
36	145	100.0	4825	10	US-09-845-416-29	Sequence 29, Appl
37	145	100.0	4825	10	US-09-845-416-29	Sequence 29, Appl
38	145	100.0	4848	10	US-09-845-416-35	Sequence 35, Appl
39	145	100.0	4966	10	US-09-845-416-28	Sequence 28, Appl
40	145	100.0	4966	10	US-09-845-416-28	Sequence 28, Appl
41	145	100.0	4390	10	US-09-845-416-34	Sequence 34, Appl
42	145	100.0	4990	10	US-09-845-416-34	Sequence 34, Appl
43	145	100.0	5060	10	US-09-845-416-36	Sequence 36, Appl
44	145	100.0	5060	10	US-09-845-416-36	Sequence 36, Appl
45	145	100.0	5149	10	US-09-845-416-27	Sequence 27, Appl

ALIGNMENTS

```

RESULT 1
US-09-782-378A-6
/ Sequence 6, Application US/09782378A
/ Patent No. US20020102731A1
/ GENERAL INFORMATION:
/ APPLICANT: Hearing, Patrick
/ APPLICANT: Bahou, Wadie
/ APPLICANT: Sandalon, Ziv
/ APPLICANT: Gnatenko, Dmitri
/ TITLE OF INVENTION: Adenoviral Vectors
/ FILE REFERENCE: STONYB-04970
/ CURRENT APPLICATION NUMBER: US/09/782,378A
/ CURRENT FILING DATE: 2001-02-12
/ PRIOR APPLICATION NUMBER: 60/237,747
/ PRIOR FILING DATE: 2000-10-02
/ NUMBER OF SEQ ID NOS: 27
/ SOFTWARE: Patentin version 3.0
/ SEQ ID NO 6
/ LENGTH: 145
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-782-378A-6

```

	Query Match	100.00%	Score 145;	DB 9;	Length 145;
	Best Local Similarity	100.00%	Pred. No. 1.6e-32;		
	Matches 145;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	TTGGCCACCTCCCTCTCTGGCGGCTCGCTCGCTCAGTCAGGCGGGGGACCAAGGTCGCC	60		
Dd	1	TTGGCCACCTCCCTCTCTGGCGGCTCGCTCGCTCAGTCAGGCGGGGGACCAAGGTCGCC	60		
Qy	61	CGACGCGCCGGGCTTTTCCCGGGCGGGCTCAGTCAGCGAGCGAGCGCGCGAGAGGGAGTG	120		
Dd	61	CGACGCGCCGGGCTTTTCCCGGGCGGGCTCAGTCAGCGAGCGAGCGCGCGAGAGGGAGTG	120		
Qy	121	GCCAACTCCATCACTAGGGGTTTCCT	145		
Dd	121	GCCAACTCCATCACTAGGGGTTTCCT	145		

```
RESULT 2
US-10-135-984-8
; Sequence 8, Application US/10135984
; Publication No. US20020182595A1
; GENERAL INFORMATION:
; APPLICANT: Matthew D. Weitzman
; APPLICANT: Anton J. Cathomen
; TITLE OF INVENTION: METHOD OF IDENTIFYING CELLULAR
; TITLE OF INVENTION: REGULATORS OF ADENO-ASSOCIATED VIRUS (AAV)
; FILE REFERENCE: SALIKINS-041A
; CURRENT APPLICATION NUMBER: US/10/135,984
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: 60/286951
; PRIOR FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 146
; TYPE: DNA
; ORGANISM: adeno-associated virus
US-10-135-984-8

Query Match          100.0%; Score 145; DB 13; Length 146;
Best Local Similarity 100.0%; Pred. No. 1.6e-32;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGGCGCTCGCTCGCTCACTAGGCGCGGCGACCAAGGTCGCC 60
DB 1 TTGGCCACTCCCTCTCTGGCGCTCGCTCGCTCACTAGGCGCGGCGACCAAGGTCGCC 60

QY 61 CGACGCCCGGGCTTTCGCCGGCGCTCAGTGAGCGGCGGCGGCGAGAGGGAGTG 120
DB 61 CGACGCCCGGGCTTTCGCCGGCGCTCAGTGAGCGGCGGCGGCGGCGAGAGGGAGTG 120

QY 121 GCCAACTCCATCACTAGGGGTTCT 145
DB 121 GCCAACTCCATCACTAGGGGTTCT 145

RESULT 3
US-09-782-378A-8
; Sequence 8, Application US/09782378A
; Patent No. US20020102731A1
; GENERAL INFORMATION:
; APPLICANT: Hearing, Patrick
; APPLICANT: Bahou, Wadie
; APPLICANT: Sandalon, Ziv
; APPLICANT: Gnatenko, Dmitri
; TITLE OF INVENTION: Adenoviral Vectors
; FILE REFERENCE: STONYB-04970
; CURRENT APPLICATION NUMBER: US/09/782,378A
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/237,747
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-378A-8

Query Match          100.0%; Score 145; DB 9; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.6e-32;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGGCGCTCGCTCGCTCACTAGGCGCGGCGACCAAGGTCGCC 60
DB 21 TTGGCCACTCCCTCTCTGGCGCTCGCTCGCTCACTAGGCGCGGCGACCAAGGTCGCC 80

QY 61 CGACGCCCGGGCTTTCGCCGGCGCTCAGTGAGCGGCGGCGGCGGCGAGAGGGAGTG 120
```

```
|||||
Db 81 CGACGCCCGGGCTTTCGCCGGCGCTCAGTGAGCGGCGGCGGCGGCGAGAGGGAGTG 140

QY 121 GCCAACTCCATCACTAGGGGTTCTCT 145
DB 141 GCCAACTCCATCACTAGGGGTTCTCT 165

RESULT 4
US-10-054-665-7
; Sequence 7, Application US/10054665
; Publication No. US20020197237A1
; GENERAL INFORMATION:
; APPLICANT: Engelhardt, John F.
; APPLICANT: Duan, Dongsheng
; TITLE OF INVENTION: Adeno-associated virus vectors
; FILE REFERENCE: 875.007US2
; CURRENT APPLICATION NUMBER: US/10/054,665
; CURRENT FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/276,625
; PRIOR FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: US 60/086,166
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: SEQ ID NO:1 of U.S. Patent No. US20020197237A1 5,478,745
US-10-054-665-7

Query Match          100.0%; Score 145; DB 13; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.6e-32;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGGCGCTCGCTCGCTCACTAGGCGCGGCGGCGACCAAGGTCGCC 60
DB 21 TTGGCCACTCCCTCTCTGGCGCTCGCTCGCTCACTAGGCGCGGCGGCGGCGACCAAGGTCGCC 80

QY 61 CGACGCCCGGGCTTTCGCCGGCGCTCAGTGAGCGGCGGCGGCGGCGGCGAGAGGGAGTG 120
DB 81 CGACGCCCGGGCTTTCGCCGGCGCTCAGTGAGCGGCGGCGGCGGCGGCGAGAGGGAGTG 140

QY 121 GCCAACTCCATCACTAGGGGTTCTCT 145
DB 141 GCCAACTCCATCACTAGGGGTTCTCT 165

RESULT 5
US-10-159-968-13/c
; Sequence 13, Application US/10159968
; Publication No. US20030152914A1
; GENERAL INFORMATION:
; APPLICANT: Musatov, Serge
; APPLICANT: Kaplitt, Michael G.
; TITLE OF INVENTION: Method for Generating Replication
; TITLE OF INVENTION: Defective Viral Vectors That are Helper Free
; FILE REFERENCE: 600-1-286
; CURRENT APPLICATION NUMBER: US/10/159,968
; CURRENT FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: US 60/294,797
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: US 60/313,007
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Adeno-associated virus
US-10-159-968-13
```

Query Match 100.0%; Score 145; DB 14; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.6e-32;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCC 60
DB 145 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCC 86

QY 61 CGAGCGCGGGCTTTGCGCGGCGGCTCGCTCACTGAGCGCGGCGAGCGCGAGAGGGAGTG 120
DB 85 CGAGCGCGGGCTTTGCGCGGCGGCTCGCTCACTGAGCGCGGCGAGCGCGAGAGGGAGTG 26

QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
DB 25 GCCAACTCCATCACTAGGGTTTCCT 1

RESULT 6
US-10-276-356-1/c
; Sequence 1, Application US/10276356
; Publication No. US20040029106A1
; GENERAL INFORMATION:
; APPLICANT: University of No. US20040029106A1th Carolina at Chapel Hill
; APPLICANT: Samulski, R. Jude
; APPLICANT: McCarty, Douglas M.
; TITLE OF INVENTION: DUPLEXED PARVOVIRUS VECTORS
; FILE REFERENCE: 5470-282
; CURRENT APPLICATION NUMBER: US/10/276,356
; CURRENT FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: PCT/US01/17587
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 175
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Inverted terminal repeat from the AAV-2 vector plasmid pSub 501
US-10-276-356-1

Query Match 100.0%; Score 145; DB 12; Length 175;
Best Local Similarity 100.0%; Pred. No. 1.6e-32;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCC 60
DB 150 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCC 91

QY 61 CGAGCGCGGGCTTTGCGCGGCGGCTCGCTCACTGAGCGCGGCGAGCGCGAGAGGGAGTG 120
DB 90 CGAGCGCGGGCTTTGCGCGGCGGCTCGCTCACTGAGCGCGGCGAGCGCGAGAGGGAGTG 31

QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
DB 30 GCCAACTCCATCACTAGGGTTTCCT 6

RESULT 7
US-10-023-208-58
; Sequence 58, Application US/10023208
; Publication No. US20030124537A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Liu, Yuan-Ching
; TITLE OF INVENTION: PROCARYOTIC LIBRARIES AND USES
; FILE REFERENCE: A-70174-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/10/023,208
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/256,163
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 63

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58
; LENGTH: 207
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic enzyme attachment site sequence
US-10-023-208-58

Query Match 100.0%; Score 145; DB 14; Length 207;
Best Local Similarity 100.0%; Pred. No. 1.5e-32;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCC 60
DB 42 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCC 101

QY 61 CGAGCGCGGGCTTTGCGCGGCGGCTCGCTCACTGAGCGCGGCGAGCGCGAGAGGGAGTG 120
DB 102 CGAGCGCGGGCTTTGCGCGGCGGCTCGCTCACTGAGCGCGGCGAGCGCGAGAGGGAGTG 161

QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
DB 162 GCCAACTCCATCACTAGGGTTTCCT 186

RESULT 8
US-09-845-416-26
; Sequence 26, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-26

Query Match 100.0%; Score 145; DB 10; Length 955;
Best Local Similarity 100.0%; Pred. No. 1.2e-32;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCC 60
DB 1 TTGGCCACTCCCTCTCTCGCGGCTCGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCC 60

QY 61 CGAGCGCGGGCTTTGCGCGGCGGCTCGCTCACTGAGCGCGGCGAGCGCGAGAGGGAGTG 120
DB 61 CGAGCGCGGGCTTTGCGCGGCGGCTCGCTCACTGAGCGCGGCGAGCGCGAGAGGGAGTG 120

QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
DB 121 GCCAACTCCATCACTAGGGTTTCCT 145

RESULT 9
US-09-845-416-26/c
; Sequence 26, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142

```
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-26

Query Match      100.0%; Score 145; DB 10; Length 955;
Best Local Similarity 100.0%; Pred. No. 1.2e-32;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAAAGTCCGC 60
Db 955 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAAAGTCCGC 896

QY 61 CGACGCCCGGGCTTTGCCGGGGGCTCAGTGAGCGGAGCGCGGCGAGAGGAGTG 120
Db 895 CGACGCCCGGGCTTTGCCGGGGGCTCAGTGAGCGGAGCGCGGCGAGAGGAGTG 836

QY 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 835 GCCAACTCCATCACTAGGGGTTCT 811

RESULT 10
US-09-845-416-33
; Sequence 33, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-33

Query Match      100.0%; Score 145; DB 10; Length 987;
Best Local Similarity 100.0%; Pred. No. 1.2e-32;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAAAGTCCGC 60
Db 955 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAAAGTCCGC 896

QY 61 CGACGCCCGGGCTTTGCCGGGGGCTCAGTGAGCGGAGCGCGGCGAGAGGAGTG 120
Db 895 CGACGCCCGGGCTTTGCCGGGGGCTCAGTGAGCGGAGCGCGGCGAGAGGAGTG 836

QY 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 835 GCCAACTCCATCACTAGGGGTTCT 811

RESULT 11
US-09-845-416-33/c
; Sequence 33, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
```

```
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-33

Query Match      100.0%; Score 145; DB 10; Length 987;
Best Local Similarity 100.0%; Pred. No. 1.2e-32;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAAAGTCCGC 60
Db 987 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAAAGTCCGC 928

QY 61 CGACGCCCGGGCTTTGCCGGGGGCTCAGTGAGCGGAGCGCGGCGAGAGGAGTG 120
Db 927 CGACGCCCGGGCTTTGCCGGGGGCTCAGTGAGCGGAGCGCGGCGAGAGGAGTG 868

QY 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 867 GCCAACTCCATCACTAGGGGTTCT 843

RESULT 12
US-09-845-416-32
; Sequence 32, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 4414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-32

Query Match      100.0%; Score 145; DB 10; Length 4414;
Best Local Similarity 100.0%; Pred. No. 9.8e-33;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAAAGTCCGC 60
Db 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAAAGTCCGC 60

QY 61 CGACGCCCGGGCTTTGCCGGGGGCTCAGTGAGCGGAGCGCGGCGAGAGGAGTG 120
Db 61 CGACGCCCGGGCTTTGCCGGGGGCTCAGTGAGCGGAGCGCGGCGAGAGGAGTG 120

QY 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 121 GCCAACTCCATCACTAGGGGTTCT 145

RESULT 13
US-09-845-416-32/c
; Sequence 32, Application US/09845416
```

```
Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 32
LENGTH: 4414
TYPE: DNA
ORGANISM: Homo sapiens
US-09-845-416-32

Query Match      100.0%; Score 145; DB 10; Length 4414;
Best Local Similarity 100.0%; Pred. No. 9.8e-33;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGGCGCGCTCGCTCAGTGGCGGGCGGCGACCAAGGTCGCC 60
Db 4414 TTGGCCACTCCCTCTCTGGCGCGCTCGCTCAGTGGCGGGCGGCGACCAAGGTCGCC 4355

QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGGCGGCGGCGGCGGCGGCGGAGTG 120
Db 4354 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGGCGGCGGCGGCGGCGGAGTG 4295

QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
Db 4294 GCCAACTCCATCACTAGGGTTTCCT 4270

RESULT 14
US-09-845-416-31
Sequence 31, Application US/09845416
Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 31
LENGTH: 4476
TYPE: DNA
ORGANISM: Homo sapiens
US-09-845-416-31

Query Match      100.0%; Score 145; DB 10; Length 4476;
Best Local Similarity 100.0%; Pred. No. 9.7e-33;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGGCGCGCTCGCTCAGTGGCGGGCGGCGACCAAGGTCGCC 60
Db 4476 TTGGCCACTCCCTCTCTGGCGCGCTCGCTCAGTGGCGGGCGGCGACCAAGGTCGCC 4417

QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGGCGGCGGCGGCGGCGGAGTG 120
Db 4416 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGGCGGCGGCGGCGGAGTG 4357

QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
Db 4356 GCCAACTCCATCACTAGGGTTTCCT 4332

Search completed: April 1, 2004, 18:35:59
Job time : 229.574 secs

Publication No. US20030171312A1
GENERAL INFORMATION:
APPLICANT: XIAO, XIAO
TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
FILE REFERENCE: DE1142
CURRENT APPLICATION NUMBER: US/09/845,416
CURRENT FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/200,777
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 31
LENGTH: 4476
TYPE: DNA
ORGANISM: Homo sapiens
US-09-845-416-31

Query Match      100.0%; Score 145; DB 10; Length 4476;
Best Local Similarity 100.0%; Pred. No. 9.7e-33;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGGCGCGCTCGCTCAGTGGCGGGCGGCGACCAAGGTCGCC 60
Db 4476 TTGGCCACTCCCTCTCTGGCGCGCTCGCTCAGTGGCGGGCGGCGACCAAGGTCGCC 4417

QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGGCGGCGGCGGCGGAGTG 120
Db 4416 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTGGCGGCGGCGGCGGAGTG 4357

QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
Db 4356 GCCAACTCCATCACTAGGGTTTCCT 4332

Search completed: April 1, 2004, 18:35:59
Job time : 229.574 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 1, 2004, 12:58:58 ; Search time 210.185 Seconds
(without alignments)
2526.461 Million cell updates/sec

Title: US-10-620-039-1_COPY_1_125
Perfect score: 125
Sequence: 1 TTGCGCACTCCCTCTCGG.....CGCAGAGAGGAGTGCGCAA 125

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001as.*
5: Geneseq2001bs.*
6: Geneseq2002s.*
7: Geneseq2003as.*
8: Geneseq2003bs.*
9: Geneseq2003cs.*
10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	125	100.0	144	8	ACF35877 AAV-2 rig
2	125	100.0	145	2	AQ41448 AAV2 Inve
3	125	100.0	145	2	AAT03385 Strict in
4	125	100.0	145	2	AAT63408 Inverted
5	125	100.0	145	2	AAX34295 Adeno-ass
6	125	100.0	145	6	ABA02990 Adeno-ass
7	125	100.0	145	6	ABS69884 Human ade
8	125	100.0	145	7	ACC58491 Adeno ass
9	125	100.0	145	8	ACF35876 AAV-2 lef
10	125	100.0	146	7	ABX93568 Adeno-ass
11	125	100.0	165	2	AQ66769 Double-D
12	125	100.0	165	2	AAT49462 Adeno ass
13	125	100.0	165	4	AAD03353 Inverted
14	125	100.0	165	6	ABS69886 Human ade
15	125	100.0	165	7	ABV77279 Nucleotid
16	125	100.0	192	2	AAT03384 Right-han
17	125	100.0	207	6	AD44821 165 bp en
18	125	100.0	955	6	AAD37254 Adeno-ass
19	125	100.0	955	6	AAD37254 Adeno-ass
20	125	100.0	987	6	AAD37261 Adeno-ass
21	125	100.0	987	6	AAD37261 Adeno-ass
22	125	100.0	4414	6	AAD37260 Adeno-ass
23	125	100.0	4414	6	AAD37260 Adeno-ass

24	125	100.0	4476	6	AAD37259
25	125	100.0	4476	6	AAD37259
26	125	100.0	4498	6	AAD37258
27	125	100.0	4498	6	AAD37258
28	125	100.0	4675	4	AAX41481
29	125	100.0	4675	4	AAF89931
30	125	100.0	4675	6	ABA02989
31	125	100.0	4675	6	ABS69879
32	125	100.0	4675	6	ABS69880
33	125	100.0	4675	9	AD676508
34	125	100.0	4679	4	AAI66974
35	125	100.0	4679	4	AAF23750
36	125	100.0	4679	6	ABK89694
37	125	100.0	4679	7	ABV76133
38	125	100.0	4680	2	AAT09008
39	125	100.0	4680	7	ABX14497
40	125	100.0	4680	8	AD50070
41	125	100.0	4683	4	AAF23749
42	125	100.0	4825	6	AAD37257
43	125	100.0	4825	6	AAD37257
44	125	100.0	4848	6	AAD37263
45	125	100.0	4848	6	AAD37263

ALIGNMENTS

RESULT 1
ACF35877
ID ACF35877 standard; DNA; 144 BP.
XX ACF35877;
XX
XX 06-NOV-2003 (first entry)
XX
XX AAV-2 right inverted terminal repeat sequence.
XX Adenovirus; Rep78; Rep68; CAP; gene therapy; AAV-2; ITR; ds.
XX Adeno associated virus.
XX WO2003061582-A2.
XX
XX 31-JUL-2003.
XX
XX 17-JAN-2003; 2003WO-US001624.
XX
XX 18-JAN-2002; 2002US-0349532P.
XX (UYDU-) UNIV DUKE.
XX
XX Li C, Zhang X;
XX WPI; 2003-627412/59.
XX
XX New recombinant adenovirus comprising an adenovirus that encodes one or more AAV REP78/68 polypeptides that are inducibly expressed, useful as a vector for gene therapy.
XX
XX Example; Page 120; 122pp; English.
XX
XX The invention relates to a recombinant adenovirus comprising an adenovirus that encodes one or more AAV REP78/68 polypeptides that are inducibly expressed. A complete virus-mediated system for recombinant AAV production is provided which comprises: (a) a first recombinant adenovirus encoding one or more AAV REP78/68 polypeptides and one or more viral capsid polypeptides; (b) a second recombinant adenovirus comprising a gene of interest and AAV inverted terminal repeats that flank the gene of interest; (c) viral helper functions; and (d) a host cell comprising the first recombinant adenovirus, the second recombinant adenovirus, and the viral helper functions. The recombinant AAV is useful as a vector for gene therapy. The present sequence represents an AAV-2 right inverted terminal repeat (ITR) sequence

```

XX Sequence 144 BP; 23 A; 49 C; 52 G; 20 T; 0 U; 0 Other;
SQ
  Query Match      100.0%; Score 125; DB 8; Length 144;
  Best Local Similarity 100.0%; Pred. No. 1.2e-23;
  Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGCGACCAAGTGC 60
Db 20 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGCGACCAAGTGC 79

Qy 61 CGACGCGCGCGGCTTTGCGCGCGGCGCTCAGTGAGCGAGCGCGGCGAGAGGAGTG 120
Db 80 CGACGCGCGGCTTTGCGCGCGGCGCTCAGTGAGCGAGCGCGGCGAGAGGAGTG 139

Qy 121 GCCAA 125
Db 140 GCCAA 144

RESULT 2
AAQ41448
ID AAQ41448 standard; DNA; 145 BP.
XX AC AAQ41448;
XX AC
XX 25-MAR-2003 (revised)
DT 27-AUG-1993 (first entry)
XX
XX AAV2 inverted terminal repeat.
XX
XX Adeno-associated virus 2; ITR; site-specific integration; vector;
XX cell-specific; gene therapy; haemoglobinopathies; thalassemia; diabetes;
XX sickle cell anaemia; cancer; parvovirus; B19; ss.
XX
XX Adeno-associated virus 2.
XX
XX Key Location/Qualifiers
XX Repeat_region 1..125 a
XX /tag= a
XX /note= "Inverted terminal repeat forming palindromic
XX hairpin"
XX misc_feature 42..83
XX /tag= b
XX /note= "Flip orientation"
XX
XX WO9309239-A1.
XX
XX 13-MAY-1993.
XX
XX 06-NOV-1992; 92WO-US009769.
XX
XX 08-NOV-1991; 91US-00789917.
XX
XX (RESE ) RESEARCH CORP TECHNOLOGIES INC.
XX
XX Srivastava A;
XX
XX WPI; 1993-167704/20.
XX
XX Adeno-associated virus-2 basal vectors - for gene therapy and treatment
XX of haemoglobinopathies and cancer etc. - has cassette contg. a promoter
XX capable of cell-specific expression, between 2 inverted terminal repeats
XX of the adeno-associated virus 2.
XX
XX Disclosure; Fig 1; 68pp; English.
XX
XX The sequence is one of two inverted terminal repeat sequences, one from
XX parvovirus B19 and the other from adeno-associated virus 2 (AAV2), used
XX in construction of an expression vector for site specific integration and
XX cell specific gene expression. The vector comprises at least one cassette
XX contg. a promoter capable of effecting cell-specific expression, operably
XX linked to a hetero- logous gene, and the cassette residing between the

```

```

CC inverted terminal repeats. The vector is safe for use in gene therapy,
CC partic. in treatment of haemoglobinopathies and a variety of diseases,
CC e.g. thalassemia, diabetes, sickle cell anaemia, and cancer. See also
CC AAQ41449. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 145 BP; 21 A; 52 C; 49 G; 23 T; 0 U; 0 Other;
  Query Match      100.0%; Score 125; DB 2; Length 145;
  Best Local Similarity 100.0%; Pred. No. 1.2e-23;
  Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGTGC 60
Db 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGTGC 60

Qy 61 CGACGCGCGCGGCTTTGCGCGCGGCGCTCAGTGAGCGAGCGCGGCGAGAGGAGTG 120
Db 61 CGACGCGCGGCTTTGCGCGCGGCGCTCAGTGAGCGAGCGCGGCGAGAGGAGTG 120

Qy 121 GCCAA 125
Db 121 GCCAA 125

RESULT 3
AAQ03385
ID AAQ03385 standard; cDNA; 145 BP.
XX AC AAT03385;
XX AC
XX 16-OCT-2003 (revised)
DT 18-APR-1996 (first entry)
XX
XX Strict inverted terminal repeat from AAV-2, used in pITRFL.
XX
XX inverted terminal repeat; ITR; adenovirus; adeno-associated virus; AAV-2;
XX replication defective; integration; gene therapy; ds.
XX
XX Adeno-associated virus 2.
XX
XX WO9523867-A1.
XX
XX 08-SEP-1995.
XX
XX 28-FEB-1995; 95WO-FR000233.
XX
XX 03-MAR-1994; 94FR-00002445.
XX
XX (RHON ) RHONE POULENC RORER SA.
XX
XX Deneffe P, Latta M, Perricaudet M, Vigne E;
XX WPI; 1995-320581/41.
XX
XX Recombinant defective adenovirus contg. integratable expression cassette
XX - for use in gene therapy to express protein, antigen or anti-sense
XX nucleic acid, also for prodn. of recombinant adeno-associated viruses.
XX
XX Example 3; Page 26; 50pp; French.
XX
XX Recombinant, non-pathogenic adenovirus which are able to integrate stably
XX into a host genome are claimed. The viruses pref. contain at least one
XX inverted terminal repeat (ITR) sequence and in particular two ITRs flank
XX a heterologous DNA insert. The present sequence is that of the strict ITR
XX from adeno associated virus AAV-2 (i.e. the ITR sequence without any
XX deletions or additions). In the construct pITRFL, the beta-galactosidase
XX marker gene is flanked by two strict AAV-2 ITRs. The defective viruses
XX are useful for stably introducing large fragments of heterologous DNA
XX making them suitable for gene therapy. (Updated on 16-OCT-2003 to
XX standardise OS field)
XX
XX Sequence 145 BP; 21 A; 52 C; 49 G; 23 T; 0 U; 0 Other;
SQ

```



```

Query Match          100.0%; Score 125; DB 2; Length 145;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGGCGGACCAAGGTGCGC 60
   |||||
Db 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGGCGGACCAAGGTGCGC 60
   |||||
QY 61 CGAGCGCGCGGCTTTGCCCGGCGGCTCACTGAGCGGAGCGGCGGAGGGAGTG 120
   |||||
Db 61 CGAGCGCGGCTTTGCCCGGCGGCTCACTGAGCGGAGCGGCGGAGGGAGTG 120
   |||||
QY 121 GCCAA 125
   |||||
Db 121 GCCAA 125
   |||||

RESULT 4
AAT63408/c
ID AAT63408 standard; DNA; 145 BP.
AC AAT63408;
XX
XX
XX
DT 18-DEC-1997 (first entry)
XX
DE Inverted terminal repeat from adeno associated virus serotype 2.
XX
XX Inverted terminal repeat; ITR; wild type; adeno associated virus; AAV;
KW serotype 2; Rep binding site; integration system; treatment; acquired;
KW inherited; disease; gene therapy; ss.
XX
XX Adeno associated virus.
XX
FH Key Location/Qualifiers
FT repeat_unit 1..42
FT /tag= a
FT /rpt_type= inverted
FT /note= "region A"
FT
FT misc_binding 20..31
FT /tag= g
FT /label= Rep_binding_site
FT repeat_unit 43..51
FT /tag= c
FT /rpt_type= inverted
FT /note= "region C"
FT
FT repeat_unit 53..61
FT /tag= d
FT /rpt_type= inverted
FT /note= "region C"
FT
FT repeat_unit 64..73
FT /tag= e
FT /rpt_type= inverted
FT /note= "region B"
FT
FT repeat_unit 75..84
FT /tag= f
FT /rpt_type= inverted
FT /note= "region B"
FT
FT repeat_unit 85..125
FT /tag= b
FT /rpt_type= inverted
FT /note= "region A"
FT
FT misc_feature 126..145
FT /tag= h
FT /note= "region D"
FT
FT
XX WO709442-A1.
XX
XX
XX 13-MAR-1997.
XX
XX 06-SEP-1996; 96WO-US014312.
XX
XX 08-SEP-1995; 95US-00525866.
XX
XX

```

```

PA (AVIG-) AVIGEN INC.
XX
XX Natsoulis G, Surosky RT;
XX
XX WPI; 1997-1992/17/17.
XX
XX Nucleic acid construct for integration into target site in recipient
PT genome - comprises target sequence to bind AAV Rep and heterologous
PT nucleotide sequence, for treatment of acquired or inherited disease.
XX
XX Claim 3; Fig 1; 54pp; English.
XX
XX The present sequence is the inverted terminal repeat (ITR) from the wild
CC type adeno associated virus (AAV) serotype 2 genome. A novel nucleic acid
CC construct, comprises a targetting sequence capable of binding AAV Rep,
CC e.g. the present sequence, and at least 1 heterologous nucleotide
CC sequence arranged relative to the targetting sequence, so that it is
CC integrated into a target site in a recipient genome. The integration
CC system can be used to integrate a selected nucleotide sequence into a
CC recipient cell genome, useful in the treatment of an acquired or
CC inherited disease, e.g. long term gene therapy. The integration system
CC has the site specific integration characteristics of AAV, is not limiting
CC in regard to the size of the nucleotide sequence to be delivered and does
CC not result in the integration of viral gene sequences into the recipient
CC cell genome. Specifically it can be used to safely and efficiently
CC deliver, and integrate large nucleotide sequences into a suitable
CC recipient cell genome without the risk of causing insertional mutagenesis
CC due to random integration events
XX
XX Sequence 145 BP; 20 A; 52 C; 49 G; 24 T; 0 U; 0 Other;
SQ
Query Match          100.0%; Score 125; DB 2; Length 145;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGGCGGACCAAGGTGCGC 60
   |||||
Db 125 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGGCGGACCAAGGTGCGC 66
   |||||
QY 61 CGAGCGCGGCTTTGCCCGGCGGCTCACTGAGCGGAGCGGCGGAGGGAGTG 120
   |||||
Db 65 CGAGCGCGGCTTTGCCCGGCGGCTCACTGAGCGGAGCGGCGGAGGGAGTG 6
   |||||
QY 121 GCCAA 125
   |||||
Db 5 GCCAA 1

RESULT 5
AAX34295
ID AAX34295 standard; DNA; 145 BP.
XX
XX AAX34295;
XX
XX 16-JUL-1999 (first entry)
XX
XX Adeno-associated virus inverted terminal repeat sequence.
DE
XX
XX Recombinant; chimeric; parvovirus; adeno-associated virus; AAV; vector;
KW promoter; rep; cap; inverted terminal repeat; ITR; erythroid cell;
KW integration; gene expression; bone marrow; peripheral blood cell;
KW endothelial cell; myocardial cell; ss.
XX
XX Adeno-associated virus.
XX
XX WO9918227-A1.
XX
XX 15-APR-1999.
XX
XX 08-OCT-1998; 98WO-US021202.
XX
XX 08-OCT-1997; 97US-0061364P.
XX
XX

```

```
PA (ADRE-) ADVANCED RES & TECHNOLOGY INST.
XX
XX Srivastava A, Ponnazhagan S;
XX
XX WPI; 1999-264033/22.
XX
XX New recombinant adeno-associated vectors.
XX
XX Claim 2; Page 69; 76pp; English.
XX
XX The invention relates to new recombinant chimeric parvovirus-adeno-
XX associated virus (AAV) vectors comprising a promoter e.g. an AAV promoter
XX (AAK34296) and a selected DNA sequence, especially an AAV rep gene and a
XX parvovirus B19 cap gene, located between 2 AAV inverted terminal repeats
XX (ITR) such as the ITR sequence shown here. The system can specifically
XX target primitive progenitor and differentiated cells of the erythroid
XX lineage, and can achieve stable integration and expression of transduced
XX genes. The vectors can be used for the in vitro or in vivo delivery of
XX genes to cells such as bone marrow cells, peripheral blood cells,
XX endothelial cells and myocardial cells
XX
XX Sequence 145 BP; 21 A; 52 C; 49 G; 23 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 125; DB 2; Length 145;
XX Best Local Similarity 100.0%; Pred. No. 1.2e-23;
XX Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGTGC 60
DB 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGTGC 60
QY 61 CGAGCGCCCGGCTTTGCCCGCGGCTCAGTGAGCGAGCGCGAGAGGAGGTG 120
DB 61 CGAGCGCCCGGCTTTGCCCGCGGCTCAGTGAGCGAGCGCGAGAGGAGGTG 120
QY 121 GCCAA 125
DB 121 GCCAA 125
XX
XX ABA02990 standard; DNA; 145 BP.
XX
XX ABA02990;
XX
XX 29-AUG-2003 (revised)
XX 19-FEB-2002 (first entry)
XX
XX Adeno-associated virus 2 ITR SEQ ID NO 2.
XX
XX Cytostatic; virucide; apoptosis inducer; p53; Saos-2 cell; infection;
XX cancer; virus; HPV16; HPV18; adeno-associated virus 2; AAV-2; ss.
XX
XX Adeno-associated virus 2.
XX
XX Key Location/Qualifiers
XX misc_structure 1..145
XX /tag= a
XX /label= ITR
XX misc_feature 72
XX /tag= b
XX /note= "unpaired base"
XX misc_feature 94
XX /tag= c
XX /note= "unpaired base"
XX
XX WO200180840-A2.
XX
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-GE001795.
XX
```

```
PR 20-APR-2000; 2000GB-00009887.
XX
XX (BTGI-) BTG INT LTD.
XX
XX Raj K, Beard PM;
XX
XX WPI; 2002-041365/05.
XX
XX Single stranded and/or looped DNA for treating mutant p53 associated
XX cancer or infection that inhibit cellular p53, having portion with an
XX base, internally located with respect to any 3' and 5' ends of the DNA.
XX
XX Example 11; Page 38; 51pp; English.
XX
XX The invention relates to single stranded and/or looped DNA having a
XX portion with at least one base, internally located with respect to any 3',
XX and 5' ends of the DNA, that is unbasepaired with another base in a form
XX that is capable of being internalised within a target cell, for use in
XX therapy, with cytostatic and virucide activity. The DNA acts as an
XX apoptosis inducer in cells that lack p53 functionality, useful for
XX killing a cell, preferably a dividing cell where the cell is other than a
XX Saos-2 cell and the DNA is not configured to express the peptide or
XX protein that selectively kills the cell. The DNA is preferably in the
XX form of an AAV or associated with AAV protein which has been treated such
XX that the DNA is no longer capable of replication or expression in cells
XX and is associated with or contained within a vehicle which is associated
XX with one or more viral fibers which facilitate internalisation of the DNA
XX into a target cell. The DNA is also useful for manufacturing a medicament
XX for treating an individual suffering from a mutant p53 associated cancer
XX or an infection that inhibits cellular p53. The method targets cancer
XX cells or cells infected with p53 inhibiting viruses, such as HPV16 or
XX HPV18 and only cells that lack p53 activity are killed and no damage to
XX cellular DNA is involved. The present sequence is that of the adeno-
XX associated virus (AAV) 2 ITR, the double loop structure found at each end
XX of the AAV DNA. (Updated on 29-AUG-2003 to standardise OS field)
XX
XX Sequence 145 BP; 24 A; 49 C; 52 G; 20 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 125; DB 6; Length 145;
XX Best Local Similarity 100.0%; Pred. No. 1.2e-23;
XX Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGTGC 60
DB 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGACCAAGTGC 80
QY 61 CGAGCGCCCGGCTTTGCCCGCGGCTCAGTGAGCGAGCGCGAGAGGAGGTG 120
DB 81 CGAGCGCCCGGCTTTGCCCGCGGCTCAGTGAGCGAGCGCGAGAGGAGGTG 140
QY 121 GCCAA 125
DB 141 GCCAA 145
XX
XX RESULT 7
XX ABS69884
XX ID ABS69884 standard; DNA; 145 BP.
XX
XX AC ABS69884;
XX
XX 21-NOV-2002 (first entry)
XX
XX Human adeno-associated virus 2 terminal repeat sequence.
XX
XX Vector; adenovirus; adeno-associated; adenosine deaminase gene; receptor;
XX adenosine deaminase deficiency; severe combined immune deficiency; PAH;
XX beta-chain; haemoglobin gene; beta-thalassaemia; sickle cell disease;
XX low density lipoprotein gene; familial hypercholesterolaemia;
XX hypoxanthine-guanine phosphoribosyltransferase; Lesch-Nyhan syndrome;
XX phenylalanine hydroxylase gene; gene therapy; phenylketonuria;
XX dystrophin gene; muscular dystrophy; cystic fibrosis; immunosuppressant;
XX human cystic fibrosis transmembrane conductance regulator gene;
```

KW antianaemic; antilipaeamic; nootropic; cytostatic; dermatological;
 XX human adeno-associated virus 2; AAV2; terminal repeat; ds.
 XX Homo sapiens.
 XX US2002102731-A1.
 XX 01-AUG-2002.
 XX 12-FEB-2001; 2001US-00782378.
 XX 02-OCT-2000; 2000US-0237747P.
 XX (UNYU) UNIV NEW YORK STATE RES FOUND.
 XX Hearing P, Bahou WF, Sandalon Z, Gnatenko DV;
 XX WPI; 2002-690619/74.
 XX Producing vector, by introducing vector having nucleotide sequence,
 PT adeno-virus inverted terminal repeats and packaging sequence, and adeno-
 PT associated virus terminal repeat, into cell, and culturing cell.
 XX Disclosure; Page 8; 191pp; English.
 XX The present invention relates to a new method of producing a vector. The
 CC method involves introducing recombinant vector having nucleotide sequence
 CC (NS) having 5' and 3' end, left and right inverted terminal repeats of
 CC adeno-virus flanking NS, adeno-virus packaging sequence linked to inverted
 CC terminal repeat, and adeno-associated virus terminal repeat linked to 3'
 CC end of NS, into cell expressing adeno-virus early gene lacking from vector
 CC ; and culturing cell to produce another vector. The method is useful for
 CC generating vectors, especially mad vectors. The method is useful in
 CC transferring nucleotide sequences of interest into a cell, for gene
 CC transfer applications (e.g. gene therapy) in vitro, ex vivo and in vivo.
 CC The nucleotide sequences are useful for treating diseases associated with
 CC it, i.e. adenosine deaminase gene associated with adenosine deaminase
 CC deficiency with severe combined immune deficiency, beta-chain of
 CC haemoglobin gene associated with beta-thalassaemia and sickle cell
 CC disease, receptor for low density lipoprotein gene associated with
 CC familial hypercholesterolaemia, hypoxanthine-guanine
 CC phosphoribosyltransferase associated with Lesch-Nyhan syndrome.
 CC Phenylalanine hydroxylase (PAH) gene associated with phenylketonuria,
 CC dystrophin gene associated with muscular dystrophy, and human cystic
 CC fibrosis transmembrane conductance regulator gene associated with cystic
 CC fibrosis. The present nucleic acid sequence represents a human adeno-
 CC associated virus 2 (AAV2) terminal repeat sequence that was used in the
 CC methods of the invention
 XX SQ Sequence 145 BP; 21 A; 52 C; 49 G; 23 T; 0 U; 0 Other;
 Query Match 100.0%; Score 125; DB 6; Length 145;
 Best Local Similarity 100.0%; Pred. No. 1.2e-23;
 Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCC 60
 Db 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCC 60
 Qy 61 CGACGCCCGGGGTTTGGCCGGCGGCTCACTGAGCGAGCGCGGCGAGAGGGAGTG 120
 Db 61 CGACGCCCGGGGTTTGGCCGGCGGCTCACTGAGCGAGCGCGGCGAGAGGGAGTG 120
 Qy 121 GCCAA 125
 Db 121 GCCAA 125
 RESULT 8
 ACC58491/c
 ID ACC58491 standard; DNA; 145 BP.
 XX
 AC ACC58491;

XX 26-AUG-2003 (first entry)
 XX Adeno associated virus inverted terminal repeat region.
 DE AAV; inverted terminal repeat; insect; gene therapy; vector; ss.
 XX Adeno associated virus.
 XX Key Location/Qualifiers
 FH stem_loop 21..145
 FT /*tag= a
 XX WO2003042361-A2.
 XX 22-MAY-2003.
 XX 08-NOV-2002; 2002WO-US035829.
 XX 09-NOV-2001; 2001US-00986618.
 PR 13-AUG-2002; 2002US-00216870.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Kotin RM, Urabe M, Ding C;
 XX WPI; 2003-449568/42.
 DR Producing an adeno-associated virus (AAV) in an insect cell, e.g. for
 PT gene therapy, comprises introducing an insect cell-compatible vector into
 PT an insect cell and maintaining the insect cell under conditions that
 PT produce AAV.
 XX Disclosure; Fig 5A; 84pp; English.
 XX The present sequence is that of an adeno associated virus (AAV)
 CC palindromic inverted terminal repeat (ITR) region. An AAV ITR is used in
 CC novel insect cell-compatible vectors of the invention that are designed
 CC for the production of AAV in insect cells. The vectors comprise an AAV
 CC ITR and coding sequences for AAV VP1, VP2 and VP3 capsid proteins, for
 CC AAV non-structural replication (Rep) proteins Rep52 or Rep40, and for AAV
 CC Rep78 or Rep68. The method of producing AAV in insect cells provides an
 CC efficient, safe and economical means of producing a large amount of
 CC recombinant AAV particles which may be used in gene therapy
 XX SQ Sequence 145 BP; 23 A; 49 C; 52 G; 21 T; 0 U; 0 Other;
 Query Match 100.0%; Score 125; DB 7; Length 145;
 Best Local Similarity 100.0%; Pred. No. 1.2e-23;
 Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCC 60
 Db 145 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTGCGCC 86
 Qy 61 CGACGCCCGGGGTTTGGCCGGCGGCTCACTGAGCGAGCGCGGCGAGAGGGAGTG 120
 Db 85 CGACGCCCGGGGTTTGGCCGGCGGCTCACTGAGCGAGCGCGGCGAGAGGGAGTG 26
 Qy 121 GCCAA 125
 Db 25 GCCAA 21
 RESULT 9
 ACF35876
 ID ACF35876 standard; DNA; 145 BP.
 XX
 AC ACF35876;
 XX 06-NOV-2003 (first entry)
 DT AAV-2 left inverted terminal repeat sequence.
 DE

```

XX Adenovirus; Rep78; Rep68; CAP; Gene therapy; AAV-2; ITR; ds.
XX Adeno associated virus.
XX WO2003061582-A2.
XX 31-JUL-2003.
XX 17-JAN-2003; 2003WO-US001624.
XX 18-JAN-2002; 2002US-0349532P.
XX (UYDU-) UNIV DUKE.
XX Li C, Zhang X;
XX WPI; 2003-627412/59.
XX New recombinant adenovirus comprising an adenovirus that encodes one or
XX more AAV REP78/68 polypeptides that are inducibly expressed, useful as a
XX vector for gene therapy.
XX Example; Page 120; 122pp; English.
XX The invention relates to a recombinant adenovirus comprising an
XX adenovirus that encodes one or more AAV REP78/68 polypeptides that are
XX inducibly expressed. A complete virus-mediated system for recombinant AAV
XX production is provided which comprises: (a) a first recombinant
XX adenovirus encoding one or more AAV REP78/68 polypeptides and one or more
XX viral capsid polypeptides; (b) a second recombinant adenovirus comprising
XX a gene of interest and AAV inverted terminal repeats that flank the gene
XX of interest; (c) viral helper functions; and (d) a host cell comprising
XX the first recombinant adenovirus, the second recombinant adenovirus, and
XX the viral helper functions. The recombinant AAV is useful as a vector for
XX gene therapy. The present sequence represents an AAV-2 left inverted
XX terminal repeat (ITR) sequence
XX Sequence 145 BP; 21 A; 52 C; 49 G; 23 T; 0 U; 0 Other;
Query Match 100.0%; Score 125; DB 8; Length 145;
Best Local Similarity 100.0%; Pred No. 1,2e-23;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGACCAAGGTCGCC 60
Db 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGACCAAGGTCGCC 60
QY 61 CGACGCCCGGGCTTTGCGCGCGCGCTCAGTGAGCGCGAGCGCGAGAGGAGTG 120
Db 61 CGACGCCCGGGCTTTGCGCGCGCGCTCAGTGAGCGCGAGCGCGAGAGGAGTG 120
QY 121 GCCAA 125
Db 121 GCCAA 125
RESULT 10
ID ABX93568
XX ABX93568 standard; DNA; 146 BP.
XX AC ABX93568;
XX DT 23-MAY-2003 (first entry)
XX DE Adeno-associated virus-2 inverted terminal repeat.
XX KW Inverted terminal repeat; Rep recognition sequence; AAV-2; ss; RRS;
XX KM viral regulatory element.
XX OS Adeno-associated virus serotype 2.
XX FH Key Location/Qualifiers

```

```

FT misc_binding 1..41
FT /tag= a
FT /bound moiety= "Nucleotides 125-85 of the present
FT sequence"
FT 42..50
FT misc_binding
FT /tag= b
FT /bound moiety= "Nucleotides 62-54 of the present
FT sequence"
FT 54..62
FT misc_binding
FT /tag= c
FT /bound moiety= "Nucleotides 50-42 of the present
FT sequence"
FT 64..72
FT misc_binding
FT /tag= d
FT /bound moiety= "Nucleotides 84-76 of the present
FT sequence"
FT 76..84
FT misc_binding
FT /tag= e
FT /bound moiety= "Nucleotides 72-64 of the present
FT sequence"
FT 85..125
FT misc_binding
FT /tag= f
FT /bound moiety= "Nucleotides 41-1 of the present sequence"
FT 93..108
FT misc_signal
FT /tag= g
FT /label= Rep recognition sequence
FT /note= "This RRS is specifically claim in claim 13"
FT 122..127
FT misc_recomb
FT /tag= h
FT /label= terminal_resolution_site
US2002182595-A1.
05-DEC-2002.
29-APR-2002; 2002US-00135984.
27-APR-2001; 2001US-0286951P.
(WEITZMAN M D.
(CATH/) CATHOMEN A J.
Weitzman MD, Cathomen AJ;
WPI; 2003-328607/31.
Identification of cellular regulations of adeno-associated virus involves
contacting an adeno-associated viral regulatory interacting molecule with
a nucleic acid comprising a adeno-associated viral regulatory element.
Disclosure; Fig 1; 16pp; English.
The invention relates to identifying molecules that interact with an
adeno-associated viral regulatory element (e.g. a Rep recognition
sequence, RRS) comprising contacting at least one molecule with a nucleic
acid comprising at least one adeno-associated viral regulatory element,
and identifying whether at least one molecule is bound to the viral
regulatory elements. Also included is a kit for analysing the interaction
between molecules and an adeno-associated viral regulatory element. The
method is useful for identifying molecules e.g. at least one cellular
protein, recombinant protein, synthetic protein or exogenous protein
(preferably human proteins that are derived from a cDNA library or
cellular lysate), or peptide, antibody, nucleic acid, lipid, carbohydrate
and/or organic or inorganic compound that interacts with an adeno-
associated viral regulatory element comprising an inverted terminal
repeat (ITR) or region. The method broadens the understanding of adeno-
associated virus (AAV) and facilitates the application of adeno-
associated virus-based vector system, therefore ITR-binding proteins and
other molecules provide new insights into the AAV life cycle, including
the regulation of gene expression and integration, and suggests
improvements for its application in therapy. The present sequence is an
AAV-2 ITR containing an RRS used in the method of the invention

```



```
FT /rpt_type= INVERTED
FT /label= B_repeat
FT 96..103
FT /*tag= g
FT /rpt_type= INVERTED
FT /label= B'_repeat
FT 104..145
FT /*tag= n
FT /rpt_type= INVERTED
FT /label= A_repeat
FT /notes= "Forms the stem of a T-shaped structure when base
FT paired with A' repeat"
FT 146..165
FT /*tag= i
FT /label= D'
FT /note= "Additional D' sequence"
XX WO9636364-A1.
XX
XX 21-NOV-1996.
XX
XX 14-MAY-1996; 96WO-USC06786.
XX
XX 15-MAY-1995; 95US-00440738.
XX
XX (SAMU/) SAMULSKI R J.
XX (XIAO/) XIAO X.
XX
XX Samulski RJ, Xiao X;
XX
XX WPI; 1997-042643/04.
XX
XX Double-D sequence directs adeno-associated virus integration into host
XX genome - used in gene therapy, maintains full length coding sequence of
XX therapeutic gene.
XX
XX Claim 1; Fig 9; 55pp; English.
XX
XX The adeno associated virus (AAV) has a 145 bp inverted terminal repeat
XX (ITR) located at each end of its genome. In addition to being able to
XX base pair with each other, the ITRs can also individually fold back on
XX themselves through the base pairing of A', B', B' and C, C' sequences
XX to form a T-shaped structure for DNA replication (see features table). It
XX has been found that viral mutants with deleted D sequences are unable to
XX replicate their DNA. Addition of a second D repeat (D') to the naturally.
XX occurring ITR resulted in a sequence, designated double-D, which was
XX sufficient to carry out the functions normally required of two wild-type
XX ITRs during a lytic AAV viral infection, i.e. it is capable of directing
XX replication and assembly into AAV, and/or the integration into the host
XX genome, of recombinant DNA containing the nucleic acid molecule. Vectors
XX and viral particles containing the double-D sequence are useful in gene
XX therapy. Replication and integration into the host genome is completely
XX effected through the double-D sequences, ensuring that the heterologous
XX gene sequences remain intact
XX
XX Sequence 165 BP; 27 A; 56 C; 56 G; 26 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 125; DB 2; Length 165;
XX Best Local Similarity 100.0%; Pred No. 1.2e-23;
XX Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 1 TTGGCCACTCCCTCTCTGCGCGTCTGCTCACTAGCGCGGCGACCAAGGTGCGC 60
Db |||||
OY 21 TTGGCCACTCCCTCTCTGCGCGTCTGCTCACTAGCGCGGCGACCAAGGTGCGC 80
Db |||||
OY 61 CGAGCGCGGCGCTTTGCGCGGCGGCTCACTAGCGCGGCGACGAGGAGGAGTG 120
Db |||||
OY 81 CGAGCGCGGCGCTTTGCGCGGCGGCTCACTAGCGCGGCGACGAGGAGGAGTG 140
Db |||||
OY 121 GCCAA 125
Db |||||
OY 141 GCCAA 145
Db |||||
```

```
RESULT 13
AAD03535
XX AAD03535 standard; DNA; 165 BP.
XX
XX AAD03535;
XX
XX 19-JUN-2001 (first entry)
XX
XX Inverted terminal repeat double DD DNA sequence.
XX
XX Recombinant adeno-associated virus; rAAV; circular intermediate; ITR;
XX inverted terminal repeat; haemostatic; antiskilling; neuroprotective;
XX antianaemic; neutropic; blood disorder; sickle cell anaemia;
XX thalassaemia; neurological disorder; haemophilia; Alzheimer's disease;
XX muscle disorder; Parkinson's disease; gene delivery; erythropoietin; epo;
XX cFTR; cystic fibrosis transmembrane conductance receptor;
XX tyrosine hydroxylase; glucocerebrosidase gene; Gaucher's disease;
XX gene therapy; Double DD; ds.
XX
XX Unidentified.
XX
XX WO200125465-A1.
XX
XX 12-APR-2001.
XX
XX 06-OCT-2000; 2000WO-US027863.
XX
XX 07-OCT-1999; 99US-0158209P.
XX
XX (IOWA ) UNIV IOWA RES FOUND.
XX (ENGE/) ENGELHARDT J F.
XX (DONG/) DONGSHENG D.
XX (ZIYI/) ZIYING Y.
XX
XX Engelhardt JF, Dongsheng D, Ziyang Y;
XX
XX WPI; 2001-266321/27.
XX
XX Composition for transferring recombinant DNAs and to express a
XX polypeptide in a host cell, comprises two recombinant adeno-associated
XX viruses.
XX
XX Disclosure; Page 139; 144pp; English.
XX
XX The patent discloses a composition comprising at least two recombinant
XX adeno-associated viruses (rAAVs). Each rAAV contains a recombinant DNA
XX molecule comprising three DNA segments linked together. The first and the
XX third DNA segments comprise 5' and 3' inverted terminal repeats (ITRs)
XX respectively from the circular intermediate of AAV. The second DNA
XX segment in each virus is different and does not comprise AAV sequence. It
XX preferably comprises sequences encoding a therapeutically effective
XX polypeptide such as the cystic fibrosis transmembrane conductance
XX receptor gene (CFTR), the erythropoietin (epo) gene, the tyrosine
XX hydroxylase gene (Parkinson's syndrome) or the glucocere- brosidase gene
XX (Gaucher's disease). The circular intermediate of AAV imparts increased
XX episomal stability and persistence of the vector in the host cell.
XX Compositions comprising rAAV sequences are useful for transferring
XX recombinant DNAs to a host cell and express a polypeptide in a host cell.
XX The recombinant vector is useful in medical therapy, which includes
XX treatment or prophylaxis of blood disorders (e.g. sickle cell anaemia,
XX thalassaemia, haemophilia), neurological disorders, such as Alzheimer's
XX disease, Parkinson's disease, muscle disorders involving skeletal,
XX cardiac or smooth muscle. AAV vector is used as a delivery vehicle for
XX gene therapy. The present sequence is the ITR DD DNA sequence referred to
XX as "double sequence". This sequence is disclosed in U.S. Patent No.
XX 5,478,745. This sequence is not used in the present invention
XX
XX Sequence 165 BP; 27 A; 56 C; 56 G; 26 T; 0 U; 0 Other;
```

```
Query Match 100.0%; Score 125; DB 4; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTGCTCACTGAGCGCGCGGACCAAGTCCGCC 60
Db 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTGCTCACTGAGCGCGCGGACCAAGTCCGCC 80
QY 61 CGACGCCCGGGCTTTGGCCCGGGCGGCTTTCAGTGAAGCGAGCGCGGAGGAGGTG 120
Db 81 CGACGCCCGGGCTTTGGCCCGGGCGGCTTTCAGTGAAGCGAGCGCGGAGGAGGTG 140
QY 121 GCCAA 125
Db 141 GCCAA 145
RESULT 14
ABS69886
ID ABS69886 standard; DNA; 165 BP.
XX AC
XX AC ABS69886;
XX DT 21-NOV-2002 (first entry)
XX DE Human adeno-associated virus 2 terminal repeat DD sequence.
XX DE Vector; adenovirus; adeno-associated; adenovirus deaminase gene; receptor;
KW adenosine deaminase deficiency; severe combined immune deficiency; PAH;
KW beta-chain; haemoglobin gene; beta-thalassaemia; sickle cell disease;
KW low density lipoprotein gene; familial hypercholesterolaemia;
KW hypoxanthine-guanine phosphoribosyltransferase; Lesch-Nyhan syndrome;
KW phenylalanine hydroxylase gene; gene therapy; phenylketonuria;
KW dystrophin gene; muscular dystrophy; cystic fibrosis; immunostimulant;
KW human cystic fibrosis transmembrane conductance regulator gene;
KW antianemic; antilipemic; nootropic; cytostatic; dermatological;
KW human adeno-associated virus 2; AAV2; terminal repeat; ds.
XX KW
XX OS Homo sapiens.
XX XX
XX US2002102731-A1.
XX PN
XX 01-AUG-2002.
XX PD
XX 12-FEB-2001; 2001US-00782378.
XX PF
XX 02-OCT-2000; 2000US-0237747P.
XX PR
XX (UYN) UNIV NEW YORK STATE RES FOUND.
XX PA
XX Hearing P, Bahou WF, Sandalon Z, Gnatenko DV;
XX PI
XX WPI; 2002-690619/74.
XX DR
XX Producing vector, by introducing vector having nucleotide sequence,
PT adenovirus inverted terminal repeats and packaging sequence, and adeno-
PT associated virus terminal repeat, into cell, and culturing cell.
XX PT
XX Disclosure; Page 8; 191pp; English.
XX PS
XX The present invention relates to a new method of producing a vector. The
CC method involves introducing recombinant vector having nucleotide sequence
CC (NS) having 5' and 3' end, left and right inverted terminal repeats of
CC adenovirus flanking NS, adenovirus packaging sequence linked to inverted
CC terminal repeat, and adeno-associated virus terminal repeat linked to 3'
CC end of NS, into cell expressing adenovirus early gene lacking from vector
CC ; and culturing cell to produce another vector. The method is useful for
CC generating vectors, especially mad vectors. The method is useful in
CC transferring nucleotide sequences of interest into a cell, for gene
CC transfer applications (e.g. gene therapy) in vitro, ex vivo and in vivo.
CC The nucleotide sequences are useful for treating diseases associated with
CC it, i.e. adenosine deaminase gene associated with adenosine deaminase
CC deficiency with severe combined immune deficiency, beta-chain of
CC haemoglobin gene associated with beta-thalassaemia and sickle cell
CC disease, receptor for low density lipoprotein gene associated with
CC familial hypercholesterolaemia, hypoxanthine-guanine

CC phosphoribosyltransferase associated with Lesch-Nyhan syndrome, ad
CC phenylalanine hydroxylase (PAH) gene associated with phenylketonuria,
CC dystrophin gene associated with muscular dystrophy, and human cystic
CC fibrosis transmembrane conductance regulator gene associated with cystic
CC fibrosis. The present nucleic acid sequence represents a human adeno-
CC associated virus 2 (AAV2) terminal repeat sequence that was used in the
CC methods of the invention
XX XX
SQ Sequence 165 BP; 24 A; 59 C; 53 G; 29 T; 0 U; 0 Other;
Query Match 100.0%; Score 125; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTGCTCACTGAGCGCGGACCAAGTCCGCC 60
Db 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTGCTCACTGAGCGCGGACCAAGTCCGCC 80
Qy 61 CGACGCCCGGGCTTTGGCCCGGGCGGCTTTCAGTGAAGCGAGCGCGGAGGAGGTG 120
Db 81 CGACGCCCGGGCTTTGGCCCGGGCGGCTTTCAGTGAAGCGAGCGCGGAGGAGGTG 140
Qy 121 GCCAA 125
Db 141 GCCAA 145
RESULT 15
ABV77279/C
ID ABV77279 standard; DNA; 165 BP.
XX AC
XX AC ABV77279;
XX DT 28-MAR-2003 (first entry)
XX DE Nucleotide sequence of pPRT, a circular adeno-associated virus.
XX KW
XX Circular adeno-associated virus; cAAV; replication; hairpin;
KW Gene therapy; pPRT; ss.
XX KW
XX Adeno associated virus.
XX OS
XX WO200297056-A2.
XX PN
XX 05-DEC-2002.
XX PD
XX 31-MAY-2002; 2002WO-US017324.
XX PF
XX 31-MAY-2001; 2001US-0294797P.
XX PR
XX 07-AUG-2001; 2001US-0313007P.
XX PR
XX (UYRQ) UNIV ROCKEFELLER.
XX PA
XX Kaplitt MG, Moussatov S;
XX PI
XX WPI; 2003-103706/09.
XX DR
XX Production of defective viral vectors for gene therapy that are
PT completely free of helper viral vectors and helper viruses.
XX PT
XX Example 4; Page 53; 69pp; English.
XX PS
XX The specification describes a nucleotide sequence capable of directing
CC circular adeno-associated virus (cAAV) replication. This nucleotide
CC sequence comprises a loop sequence (TGGCCAA) flanked on the 5' and 3'
CC sides by complementary sequences. The cAAVs are useful in gene
CC between the complementary sequences. The cAAVs are useful in gene
CC therapy, e.g. to treat an acute medical condition a nucleic acid encoding
CC a therapeutic protein is inserted into the cAAV. The present sequence
CC represents pPRT, a cAAV that contains a wild-type circularization point
CC (the PRT domain), consisting of a single ITR flanked by two D-sequences
XX Sequence 165 BP; 26 A; 56 C; 56 G; 27 T; 0 U; 0 Other;
XX SQ

Query Match 100.0%; Score 125; DB 7; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGCGACCAAGGTGCCC	60
DB	145	TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGCGACCAAGGTGCCC	86
QY	61	CGACGCGCGCGGCTTTGGCCCGCGCGCTCAGTGAGCGCGCGCGAGCGCGAGAGGGAGTG	120
DB	85	CGACGCGCGCGGCTTTGGCCCGCGCGCTCAGTGAGCGCGCGAGCGCGAGAGGGAGTG	26
QY	121	GCCAA 125	
DB	25	GCCAA 21	

Search completed: April 1, 2004, 15:07:23
Job time : 210.185 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 1, 2004, 16:04:14 ; Search time 188.426 Seconds

(without alignments)

2480.895 Million cell updates/sec

Title: US-10-620-039-1_COPY_1_125

Perfect score: 125

Sequence: 1 TTGGCCACTCCCTCTCTGCG.....CGCAGAGAGGAGTGCGCAA 125

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2465228 seqs, 1869859620 residues

Total number of hits satisfying chosen parameters: 4930456

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA:*

1: /cgn2_6/ptodata/1/pubpna/PCT PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT NEW PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09_NEW PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW PUB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW PUB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	125	100.0	130	9	US-09-928-158B-1
2	125	100.0	145	9	US-09-782-378A-6
3	125	100.0	145	14	US-10-240-198-2
4	125	100.0	146	13	US-10-135-984-8
5	125	100.0	165	9	US-09-782-378A-8
6	125	100.0	165	13	US-10-054-665-7
7	125	100.0	165	14	US-10-159-968-13
8	125	100.0	175	12	US-10-276-356-1
9	125	100.0	207	14	US-10-023-208-58
10	125	100.0	955	10	US-09-845-416-26
11	125	100.0	955	10	US-09-845-416-26
12	125	100.0	987	10	US-09-845-416-33
13	125	100.0	987	10	US-09-845-416-33
14	125	100.0	4414	10	US-09-845-416-32
15	125	100.0	4414	10	US-09-845-416-32

c 16	125	100.0	4476	10	US-09-845-416-31	Sequence 31, Appl
17	125	100.0	4476	10	US-09-845-416-31	Sequence 31, Appl
18	125	100.0	4498	10	US-09-845-416-30	Sequence 30, Appl
c 19	125	100.0	4498	10	US-09-845-416-30	Sequence 30, Appl
20	125	100.0	4675	9	US-09-782-378A-1	Sequence 1, Appl
21	125	100.0	4675	9	US-09-782-378A-2	Sequence 2, Appl
22	125	100.0	4675	14	US-10-240-198-1	Sequence 1, Appl
23	125	100.0	4675	14	US-10-291-583-7	Sequence 7, Appl
24	125	100.0	4675	9	US-09-804-898-1	Sequence 1, Appl
25	125	100.0	4679	9	US-09-945-681-10	Sequence 10, Appl
26	125	100.0	4679	13	US-10-038-972A-12	Sequence 12, Appl
27	125	100.0	4679	14	US-10-136-819-6	Sequence 6, Appl
28	125	100.0	4680	13	US-10-077-294-1	Sequence 1, Appl
29	125	100.0	4680	13	US-10-163-886-1	Sequence 1, Appl
30	125	100.0	4680	14	US-10-263-127-1	Sequence 1, Appl
31	125	100.0	4680	14	US-10-375-777-1	Sequence 1, Appl
32	125	100.0	4681	12	US-10-696-261-18	Sequence 18, Appl
33	125	100.0	4681	12	US-10-696-282-18	Sequence 18, Appl
34	125	100.0	4681	12	US-10-696-900-18	Sequence 18, Appl
35	125	100.0	4683	12	US-10-696-261-19	Sequence 19, Appl
36	125	100.0	4683	12	US-10-696-282-19	Sequence 19, Appl
37	125	100.0	4683	12	US-10-696-900-19	Sequence 19, Appl
38	125	100.0	4825	10	US-09-845-416-29	Sequence 29, Appl
39	125	100.0	4825	10	US-09-845-416-29	Sequence 29, Appl
c 40	125	100.0	4848	10	US-09-845-416-35	Sequence 35, Appl
c 41	125	100.0	4848	10	US-09-845-416-35	Sequence 35, Appl
c 42	125	100.0	4866	10	US-09-845-416-28	Sequence 28, Appl
c 43	125	100.0	4866	10	US-09-845-416-28	Sequence 28, Appl
c 44	125	100.0	4990	10	US-09-845-416-34	Sequence 34, Appl
c 45	125	100.0	4990	10	US-09-845-416-34	Sequence 34, Appl
	125	100.0	5060	10	US-09-845-416-36	Sequence 36, Appl

ALIGNMENTS

RESULT 1
US-09-928-158B-1
; Sequence 1, Application US/09928158B
; Patent No. US2002017722A1
; GENERAL INFORMATION:
; APPLICANT: SIKUN, LI
; TITLE OF INVENTION: REPLICATION COMPETENT AAV HELPER FUNCTIONS
; FILE REFERENCE: 102182-18
; CURRENT APPLICATION NUMBER: US/09/928,158B
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/224,132
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
; LENGTH: 130
; TYPE: DNA
; ORGANISM: adeno-associated virus 2
US-09-928-158B-1

Query Match 100.0%; Score 125; DB 9; Length 130;
Best Local Similarity 100.0%; Pred. No. 1.4e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGCGCAAGGTCGCC	60
Db	1	TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGCGCAAGGTCGCC	60
QY	61	CGACGCCCGGGCTTTGCGCGCGCGCTCACTGAGCGCGCGCGCGAGGAGGAGTG	120
Db	61	CGACGCCCGGGCTTTGCGCGCGCGCTCACTGAGCGCGCGCGCGAGGAGGAGTG	120
QY	121	GCCTAA 125	
Db	121	GCCTAA 125	

RESULT 2

```

US-09-782-378A-6
; Sequence 6, Application US/09782378A
; Patent No. US20020102731A1
; GENERAL INFORMATION:
; APPLICANT: Hearing, Patrick
; APPLICANT: Bahou, Wadie
; APPLICANT: Sadalon, Ziv
; APPLICANT: Gnatenko, Dmitri
; TITLE OF INVENTION: Adenoviral Vectors
; FILE REFERENCE: STONYB-04970
; CURRENT APPLICATION NUMBER: US/09/782,378A
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/237,747
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 145
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-378A-6

```

Query Match	100.0%;	Score 125;	DB 9;	Length 145;
Best Local Similarity	100.0%;	Pred. No. 1.3e-25;		
Matches 125;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	TTGGGCACCTCTCTCGCGGCTCGCTCGCTACTGAGCGCGGCGACCAAGGTCGCC	60	
Db	1	TTGGGCACCTCTCTCTCGCGGCTCGCTCGCTACTGAGCGCGGCGACCAAGGTCGCC	60	
Qy	61	CGAGCGCCCGGGCTTTTGGCCCGGCGGCTTCAGTGAGCGAGCGCGCGAGAGGGAGTG	120	
Db	61	CGAGCGCCCGGGCTTTTGGCCCGGCGGCTTCAGTGAGCGAGCGCGCGAGAGGGAGTG	120	
Qy	121	GCCAA	125	
Db	121	GCCAA	125	

```

RESULT 3
US-10-240-198-2
Sequence 2, Application US/10240198
Publication No. US20030100115A1
GENERAL INFORMATION:
APPLICANT: BTG International Ltd
APPLICANT: BEARD DR, PETER
APPLICANT: RAJ DR, KENNETH
TITLE OF INVENTION: CYTOTOXIC AGENTS
FILE REFERENCE: 142184WO
CURRENT APPLICATION NUMBER: US/10/240,198
CURRENT FILING DATE: 2002-09-30
PRIOR APPLICATION NUMBER: 000987.1
PRIOR FILING DATE: 2000-04-20
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 145
TYPE: DNA
ORGANISM: adeno-associated virus 2
FEATURE:
NAME/KEY: misc_structure
LOCATION: (1)..(145)
OTHER INFORMATION: ITR
FEATURE:
NAME/KEY: misc_feature
LOCATION: (72)
OTHER INFORMATION: Unpaired base
FEATURE:
NAME/KEY: misc_feature
LOCATION: (94)
OTHER INFORMATION: Unpaired base
US-10-240-198-2

```

```

Query Match          100.0%; Score 125; DB 14; Length 145;
Best Local Similarity 100.0%; Pred. NO. 1.3e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY      1  TTGGGCACATCCCTCTCTGCGGCTCGCTCAGTCAGGCGCGGGCGACCAAGGTCGCC 60
DB      21  TTGGGCACATCCCTCTCTGCGGCTCGCTCAGTCAGGCGCGGGCGACCAAGGTCGCC 80

QY      61  CGAGCCCGGGCTTTGCCC CGGCGGCTCAGTCAGGCGAGCGAGCGCGCAGAGGGAGTG 120
DB      81  CGAGCCCGGGCTTTGCCC CGGCGGCGCTCAGTCAGGCGAGCGAGCGCGCGCAGAGGGAGTG 140

QY      121 GCCAA 125
DB      141 GCCAA 145

RESULT 4
US-10-135-984-8
; Sequence 8, Application US/10135984
; Publication NO. US20020182595A1
; GENERAL INFORMATION:
; APPLICANT: Matthew D. Weitzman
; APPLICANT: Anton J. Cathomen
; TITLE OF INVENTION: METHOD OF IDENTIFYING CELLULAR
; TITLE OF INVENTION: REGULATORS OF ADENO-ASSOCIATED VIRUS (AAV)
; FILE REFERENCE: SALKINS.041A
; CURRENT APPLICATION NUMBER: US/10/135,984
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: 60/286951
; PRIOR FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 146
; TYPE: DNA
; ORGANISM: adeno-associated virus
US-10-135-984-8

```

```

Query Match          100.0%; Score 125; DB 13; Length 146;
Best Local Similarity 100.0%; Pred.No. 1.3e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy      1  TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTGAGCGCGGGCGAGCCAAAGGTCGCC 60
Db      1  TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTGAGCGCGGGCGAGCCAAAGGTCGCC 60

Qy      61  CGACGCCCGGGCTTTGCCGGCGCGCTCACTGAGCGAGCGAGCGCGAGAGGGAGTG 120
Db      61  CGACGCCCGGGCTTTGCCGGCGCGCTCACTGAGCGAGCGAGCGCGAGAGGGAGTG 120

Qy      121  GCCAA 125
Db      121  GCCAA 125

RESULT 5
US-09-782-378A-8
; Sequence 8, Application US/09782378A
; Patent No. US20020102731A1
; GENERAL INFORMATION:
; APPLICANT: Hearing, Patrick
; APPLICANT: Bahou, Wadie
; APPLICANT: Sandalon, Ziv
; APPLICANT: Gnatenko, Dmitri
; TITLE OF INVENTION: Adenoviral Vectors
; FILE REFERENCE: STONYB-04970
; CURRENT APPLICATION NUMBER: US/09/782,378A
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/237,747
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0

```

```

; SEQ ID NO 8
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-378A-8

Query Match      100.0%; Score 125; DB 9; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.3e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTAGAGCGCGGCGACCAAGGTGCGC 60
DB 21 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTAGAGCGCGGCGACCAAGGTGCGC 80

QY 61 CGAGCGCGGCTTTGCCCGGCGCTCACTAGAGCGAGCGAGCGCGAGAGGAGTG 120
DB 81 CGAGCGCGGCTTTGCCCGGCGCTCACTAGAGCGAGCGAGCGCGAGAGGAGTG 140

QY 121 GCCAA 125
DB 141 GCCAA 145

RESULT 6
US-10-054-665-7
; Sequence 7, Application US/10054665
; Publication No. US20020197237A1
; GENERAL INFORMATION:
; APPLICANT: Engelhardt, John F.
; APPLICANT: Duan, Dongsheng
; TITLE OF INVENTION: Adeno-associated virus vectors
; FILE REFERENCE: 875.007US2
; CURRENT APPLICATION NUMBER: US/10/054,665
; PRIOR FILING DATE: 2002-06-13
; PRIOR FILING DATE: 1999-03-25
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: SEQ ID NO:1 of U.S. Patent No. US20020197237A1 5,478,745
US-10-054-665-7

Query Match      100.0%; Score 125; DB 13; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.3e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTAGAGCGCGGCGACCAAGGTGCGC 60
DB 21 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTAGAGCGCGGCGACCAAGGTGCGC 80

QY 61 CGAGCGCGGCTTTGCCCGGCGCTCACTAGAGCGAGCGAGCGCGAGAGGAGTG 120
DB 81 CGAGCGCGGCTTTGCCCGGCGCTCACTAGAGCGAGCGAGCGCGAGAGGAGTG 140

QY 121 GCCAA 125
DB 141 GCCAA 145

RESULT 7
US-10-159-968-13/C
; Sequence 13, Application US/10159968
; Publication No. US20030152914A1
; GENERAL INFORMATION:
; APPLICANT: Kaplitt, Michael G.
; APPLICANT: Musatov, Serge
; TITLE OF INVENTION: Method for Generating Replication

```

```

; TITLE OF INVENTION: Defective Viral Vectors That are Helper Free
; FILE REFERENCE: 600-1-286
; CURRENT APPLICATION NUMBER: US/10/159,968
; CURRENT FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: US 60/294,797
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: US 60/313,007
; PRIOR FILING DATE: 2001-08-07
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Adeno-associated virus
US-10-159-968-13

Query Match      100.0%; Score 125; DB 14; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.3e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTAGAGCGCGGCGACCAAGGTGCGC 60
DB 145 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTAGAGCGCGGCGACCAAGGTGCGC 86

QY 61 CGAGCGCGGCTTTGCCCGGCGCTCACTAGAGCGAGCGAGCGCGAGAGGAGTG 120
DB 85 CGAGCGCGGCTTTGCCCGGCGCTCACTAGAGCGAGCGAGCGCGAGAGGAGTG 26

QY 121 GCCAA 125
DB 25 GCCAA 21

RESULT 8
US-10-276-356-1/C
; Sequence 1, Application US/10276356
; Publication No. US20040029106A1
; GENERAL INFORMATION:
; APPLICANT: University of No. US20040029106A1h Carolina at Chapel Hill
; APPLICANT: Samulski, R. Jude
; APPLICANT: McCarthy, Douglas M.
; TITLE OF INVENTION: DUPLEXED PARVOVIRUS VECTORS
; FILE REFERENCE: 5470-282
; CURRENT APPLICATION NUMBER: US/10/276,356
; CURRENT FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: PCT/US01/17587
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 175
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Inverted terminal repeat from the AAV-2 vector plasmid pSub 201
US-10-276-356-1

Query Match      100.0%; Score 125; DB 12; Length 175;
Best Local Similarity 100.0%; Pred. No. 1.3e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTAGAGCGCGGCGACCAAGGTGCGC 60
DB 150 TTGGCCACTCCCTCTCTGCGGCTCGCTCACTAGAGCGCGGCGACCAAGGTGCGC 91

QY 61 CGAGCGCGGCTTTGCCCGGCGCTCACTAGAGCGAGCGAGCGCGAGAGGAGTG 120
DB 90 CGAGCGCGGCTTTGCCCGGCGCTCACTAGAGCGAGCGAGCGCGAGAGGAGTG 31

QY 121 GCCAA 125
DB 30 GCCAA 26

```

```
RESULT 9
US-10-023-208-58
; Sequence 58, Application US/10023208
; Publication No. US20030124537A1
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; APPLICANT: Liu, Yuan-Ching
; TITLE OF INVENTION: PROCAIOTIC LIBRARIES AND USES
; FILE REFERENCE: A-70174-1/AFT/RMS/RVK
; CURRENT APPLICATION NUMBER: US/10/023,208
; CURRENT FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/256,163
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 58
; LENGTH: 207
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic enzyme attachment site sequence
US-10-023-208-58

Query Match      100.0%; Score 125; DB 14; Length 207;
Best Local Similarity 100.0%; Pred. No. 1.2e-25;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGCGACCAAGTGC 60
Db 42 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGCGACCAAGTGC 101
Qy 61 CGACGCCCGGGCTTTGCCCGCGCGCTCACTGAGCGCGCGGCGAGAGGAGTG 120
Db 102 CGACGCCCGGGCTTTGCCCGCGCGCTCACTGAGCGCGCGGCGAGAGGAGTG 161
Qy 121 GCCAA 125
Db 162 GCCAA 166

RESULT 10
US-09-845-416-26
; Sequence 26, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-26

Query Match      100.0%; Score 125; DB 10; Length 955;
Best Local Similarity 100.0%; Pred. No. 8.9e-26;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGCGACCAAGTGC 60
Db 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGCGACCAAGTGC 60
Qy 61 CGACGCCCGGGCTTTGCCCGCGCGCTCACTGAGCGCGCGGCGAGAGGAGTG 120
Db 61 CGACGCCCGGGCTTTGCCCGCGCGCTCACTGAGCGCGCGGCGAGAGGAGTG 120
```

```
Qy 121 GCCAA 125
Db 121 GCCAA 125

RESULT 11
US-09-845-416-26/c
; Sequence 26, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-26

Query Match      100.0%; Score 125; DB 10; Length 955;
Best Local Similarity 100.0%; Pred. No. 8.9e-26;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGCGACCAAGTGC 60
Db 955 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGCGACCAAGTGC 896
Qy 61 CGACGCCCGGGCTTTGCCCGCGCGCTCACTGAGCGCGCGGCGAGAGGAGTG 120
Db 895 CGACGCCCGGGCTTTGCCCGCGCGCTCACTGAGCGCGCGGCGAGAGGAGTG 836
Qy 121 GCCAA 125
Db 835 GCCAA 831

RESULT 12
US-09-845-416-33
; Sequence 33, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-33

Query Match      100.0%; Score 125; DB 10; Length 987;
Best Local Similarity 100.0%; Pred. No. 8.8e-26;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGCGACCAAGTGC 60
Db 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGCGACCAAGTGC 60
```

Qy	61	CGACGCCCGGGCTTTGCCCGCGGCTTCAGTGAAGCGAGCGCGCAGAGGGAGTG	120
Db	61	CGACGCCCGGGCTTTGCCCGCGGCTTCAGTGAAGCGAGCGCGCAGAGGGAGTG	120
Qy	121	GCCAA	125
Db	121	GCCAA	125

```

RESULT 13
US-09-845-415-33/c
; Sequence 33, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-415-33

```

	Query Match	100.0%	Score 125;	DB 10;	Length 987;
	Best Local Similarity	100.0%;	Pred. No. 8.8e-26;		
	Matches 125;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	TTGGCCACTCCCTCTCTGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC	60		
Db	987				
QY	61	CGACGCCCGGGCTTTGCCCGGCGGCTCACTGAGCGAGCGCGCAGAGGGAGTG	120		
Db	927	CGACGCCCGGGCTTTGCCCGGCGGCTCACTGAGCGAGCGCGCAGAGGGAGTG	868		
QY	121	GCCAA 125			
Db	867	GCCAA 863			

```

RESULT 14
US-09-845-416-32
; Sequence 32, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 32
; LENGTH: 4414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-32

```

```

Query Match      100.0%; Score 125; DB 10; Length 4414;
Best Local Similarity 100.0%; Pred. No. 6.3e-26;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGGCACCTCCCTCTCTCGGCGCTGCTCGCTACTAGCGCGGCGGACCAAGGTCCGC 60

```

Db	1	TTGGCACATCCCTCTCTGGCGCTCGCTCGCTCACTGAGCGGGCGACCAAAAGGTGCGCC	60
Qy	61	CGACGCCCGGGCTTTGCGCGGGCGGCGCTCAGTGAGCGAGCGAGCGCGCAGAGGGAGTG	120
Db	61	CGACGCCCGGGCTTTGCGCGGGCGGCGCTCAGTGAGCGAGCGAGCGCGCAGAGGGAGTG	120
Qy	121	GCCAA 125	
Db	121	GCCAA 125	

```

RESULT 15
US-09-845-416-32/C
; Sequence 32, Application US/09845416
; Publication No. US20030171312A1
; GENERAL INFORMATION:
; APPLICANT: XIAO, XIAO
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A DYSTROPHY MINIGENE AND USE
; FILE REFERENCE: DE1142
; CURRENT APPLICATION NUMBER: US/09/845,416
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,777
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 4414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-845-416-32

```

	Query Match.	100.0%;	Score 125;	DB 10;	Length 4414;
	Best Local Similarity	100.0%;	Pred. No. 6.3e-26;		
	Matches 125;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	TTGGCCACTCCCTCTCTGGCGGCTCGCTCGCTCACTGAGCGCGGGCGACCAAGGTCGCC	60		
Db	4414	TTGGCCACTCCCTCTCTGGCGGCTCGCTCGCTCACTGAGCGCGGGCGACCAAGGTCGCC	4355		
Qy	61	CGAGCGCCGGGCTTTTGCCCGGGGGGGCTCAGTGAGCGGAGCGCGCGCAGAGAGGGAGTG	120		
Db	4354	CGAGCGCCGGGCTTTTGCCCGGGGGGGCTCAGTGAGCGGAGCGCGCGCAGAGAGGGAGTG	4295		
Qy	121	GCCAA 125			
Db	4294	GCCAA 4290			

Search completed: April 1, 2004, 18:36:00
Job time : 189.426 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 1, 2004, 14:52:43 ; Search time 61.7593 Seconds
(without alignments)
1302.928 Million cell updates/sec

Title: US-10-620-039-1

Perfect score: 145

Sequence: 1 TTGGCCACTCCTCTCTGCGG.....CTCCATCACTAGGGGTTTCCT 145

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/prodata/2/ina/5A COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCITUS COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	145	100.0	145	1	US-07-789-917A-1
2	145	100.0	145	3	US-08-702-573-4
3	145	100.0	145	3	US-07-982-193-1
4	145	100.0	145	1	US-07-989-841A-1
5	145	100.0	145	2	US-08-440-738A-1
6	145	100.0	145	3	US-08-471-914-1
7	145	100.0	145	4	US-09-276-625-7
8	145	100.0	4680	1	US-08-254-358-1
9	145	100.0	4680	1	US-08-475-391-1
10	145	100.0	4680	2	US-08-709-609-1
11	145	100.0	4680	5	PCT-US95-07178-1
12	145	100.0	5932	4	US-09-299-141-4
13	145	100.0	5932	4	US-09-299-141-4
14	145	100.0	6142	4	US-09-299-141-8
15	145	100.0	6142	4	US-09-299-141-8
16	145	100.0	6253	3	US-08-893-327-15
17	145	100.0	6253	3	US-08-893-327-15
18	145	100.0	6280	3	US-08-893-327-17
19	145	100.0	6280	3	US-08-893-327-17
20	145	100.0	6280	3	US-08-893-327-19
21	145	100.0	6280	3	US-08-893-327-19
22	145	100.0	6565	4	US-09-299-141-1
23	145	100.0	6565	4	US-09-299-141-1
24	145	100.0	6714	4	US-09-299-141-6
25	145	100.0	6714	4	US-09-299-141-6
26	145	100.0	6924	4	US-09-299-141-9
27	145	100.0	6924	4	US-09-299-141-9

Sequence 10, Appl
Sequence 10, Appl
Sequence 11, Appl
Sequence 11, Appl
Sequence 7, Appl
Sequence 3, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 5, Appl
Sequence 5, Appl
Sequence 2, Appl
Sequence 4, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 6, Appl
Sequence 6, Appl

ALIGNMENTS

RESULT 1

US-07-789-917A-1

; Sequence 1, Application US/07789917A

; Patent No. 5252479

; GENERAL INFORMATION:

; APPLICANT: Srivastava, Arun

; TITLE OF INVENTION: SAFE VECTOR FOR GENE THERAPY

; NUMBER OF SEQUENCES: 2

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Scully, Scott, Murphy Presser

; STREET: 400 Garden City Plaza

; CITY: Garden City

; STATE: New York

; COUNTRY: USA

; ZIP: 11530

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release q.0, Version q.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07789,917A

; FILING DATE: 19911118

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: McNulty, William E.

; REGISTRATION NUMBER: 22,606

; REFERENCE/DOCKET NUMBER: 8361

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (516) 742-4343

; TELEFAX: (516) 742-4366

; TELEX: 230 901 SANS UR

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 145 base pairs

; TYPE: NUCLEIC ACID

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-07-789-917A-1

Query Match 100.0%; Score 145; DB 1; Length 145;

Best Local Similarity 100.0%; Pred. No. 4.3e-31;

Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCTCTCTGCGGCTCGCTCACTAGGCGCGCGCCACCAAGGTGCGC 60

Db 1 TTGGCCACTCCTCTCTGCGGCTCGCTCACTAGGCGCGCGCCACCAAGGTGCGC 60

QY 61 CGACGCCCGGCGCTTTGCCCGGCGCGCTCACTAGGCGCGCGCCACCAAGGTG 120

Db 61 CGACGCCCGGCTTCCCGCGGGCTCAGTGAGCGAGCGCGCAGAGGAGTG 120
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 121 GCCAACTCCATCACTAGGGGTTCT 145

RESULT 2
US-08-702-573-4
; Sequence 4, Application US/08702573
; Patent No. 6033885
; GENERAL INFORMATION:
; APPLICANT: LATTA, Martine
; APPLICANT: DENEPIE, Patrice
; APPLICANT: VIGNE, Emmanuelle
; APPLICANT: PERRICAUDET, Michel
; TITLE OF INVENTION: INTEGRATIVE RECOMBINANT ADENOVIRUSES,
; TITLE OF INVENTION: PREPARATION THEREOF AND THERAPEUTICAL USES THEREOF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd. 343
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,573
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94/02445
; FILING DATE: 03-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/FR95/00233
; FILING DATE: 28-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: 38,619
; REFERENCE/DOCKET NUMBER: ST94011-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)454-3839
; TELEFAX: (610)454-3808
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..145
; OTHER INFORMATION: /note= "Minimal ITR Sequence"
US-08-702-573-4

Query Match 100.0%; Score 145; DB 3; Length 145;
Best Local Similarity 100.0%; Pred. No. 4.3e-31;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
Db 1 TTGGCCACTCCCTCTCGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
QY 61 CGACGCCCGGCTTTCGCCGGCGGCTCAGTGAGCGAGCGCGCAGAGGAGTG 120
Db 61 CGACGCCCGGCTTTCGCCGGCGGCTCAGTGAGCGAGCGCGCAGAGGAGTG 120

QY 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 121 GCCAACTCCATCACTAGGGGTTCT 145
RESULT 3
US-07-982-193-1
; Sequence 1, Application US/07982193
; Patent No. 6261834
; GENERAL INFORMATION:
; APPLICANT: Srivastava, Arun
; TITLE OF INVENTION: SAFE VECTOR FOR GENE THERAPY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/982,193
; FILING DATE: 19921125
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McNulty, William E.
; REGISTRATION NUMBER: 22,606
; REFERENCE/DOCKET NUMBER: 8361
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 145 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-982-193-1
Query Match 100.0%; Score 145; DB 3; Length 145;
Best Local Similarity 100.0%; Pred. No. 4.3e-31;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTCGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
Db 1 TTGGCCACTCCCTCTCTCGCGGCTCGCTCACTGAGCGCGGCGACCAAGGTCGCC 60
QY 61 CGACGCCCGGCTTTCGCCGGCGGCTCAGTGAGCGAGCGCGCAGAGGAGTG 120
Db 61 CGACGCCCGGCTTTCGCCGGCGGCTCAGTGAGCGAGCGCGCAGAGGAGTG 120
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 121 GCCAACTCCATCACTAGGGGTTCT 145
RESULT 4
US-07-989-841A-1
; Sequence 1, Application US/07989841A
; Patent No. 5478745
; GENERAL INFORMATION:
; APPLICANT: Samulski, R. J.
; APPLICANT: Xiao, X.
; TITLE OF INVENTION: Recombinant Viral Vector System
; NUMBER OF SEQUENCES: 6

```
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/989,841A
; FILING DATE: On even date herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 6636-013
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 165 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; US-07-989-841A-1

Query Match      100.0%; Score 145; DB 1; Length 165;
Best Local Similarity 100.0%; Pred. No. 4.3e-31;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGGACCAAGGTGCGCC 60
DB      21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGGACCAAGGTGCGCC 80

QY      61 CGACGCCCGGGCTTTGCCCGGGCGGCTCGCTCACTGAGCGAGCGCGGCGGAGGAGTG 120
DB      81 CGACGCCCGGGCTTTGCCCGGGCGGCTCGCTCACTGAGCGAGCGCGGCGGAGGAGTG 140

QY      121 GCCAACTCCATCACTAGGGTTCTCT 145
DB      141 GCCAACTCCATCACTAGGGTTCTCT 165

RESULT 5
US-08-440-738A-1
; Sequence 1, Application US/08440738A
; Patent No. 5869305
; GENERAL INFORMATION:
; APPLICANT: Samulski, R. J.
; APPLICANT: Xiao, X.
; TITLE OF INVENTION: Recombinant Viral Vector System
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,738A

Query Match      100.0%; Score 145; DB 1; Length 165;
Best Local Similarity 100.0%; Pred. No. 4.3e-31;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGGACCAAGGTGCGCC 60
DB      21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGGACCAAGGTGCGCC 80

QY      61 CGACGCCCGGGCTTTGCCCGGGCGGCTCGCTCACTGAGCGAGCGCGGCGGAGGAGTG 120
DB      81 CGACGCCCGGGCTTTGCCCGGGCGGCTCGCTCACTGAGCGAGCGCGGCGGAGGAGTG 140

QY      121 GCCAACTCCATCACTAGGGTTCTCT 145
DB      141 GCCAACTCCATCACTAGGGTTCTCT 165

RESULT 6
US-08-471-914-1
; Sequence 1, Application US/08471914A
; Patent No. 6057152
; GENERAL INFORMATION:
; APPLICANT: Samulski, R.
; APPLICANT: Xiao, X.
; TITLE OF INVENTION: RECOMBINANT VIRAL VECTOR SYSTEM
; FILE REFERENCE: 6636-027
; CURRENT APPLICATION NUMBER: US/08/471,914A
; CURRENT FILING DATE: 1995-06-06
; EARLIER APPLICATION NUMBER: 08/440,738
; EARLIER FILING DATE: 1995-05-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: double-D
; OTHER INFORMATION: sequence
; US-08-471-914-1

Query Match      100.0%; Score 145; DB 3; Length 165;
Best Local Similarity 100.0%; Pred. No. 4.3e-31;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGGACCAAGGTGCGCC 60
DB      21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGGACCAAGGTGCGCC 80

QY      61 CGACGCCCGGGCTTTGCCCGGGCGGCTCGCTCACTGAGCGAGCGCGGCGGAGGAGTG 120
DB      81 CGACGCCCGGGCTTTGCCCGGGCGGCTCGCTCACTGAGCGAGCGCGGCGGAGGAGTG 140

QY      121 GCCAACTCCATCACTAGGGTTCTCT 145
```



```
Db 141 GCCAACTCCATCACTAGGGTTCTT 165
|||||
RESULT 7
US-09-276-625-7
; Sequence 7, Application US/09276625
; Patent No. 6436392
; GENERAL INFORMATION:
; APPLICANT: Engelhardt, John F.
; TITLE OF INVENTION: Adeno-associated virus vectors
; FILE REFERENCE: 875.007U1
; CURRENT APPLICATION NUMBER: US/09/276,625
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/086,166
; PRIOR FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 165
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: SEQ ID NO:1 of U.S. Patent No. 6436392 5,478,745
US-09-276-625-7
Query Match 100.0%; Score 145; DB 4; Length 165;
Best Local Similarity 100.0%; Pred. No. 4.3e-31;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGCGACCAAGGTGCGCC 60
Db 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGCGACCAAGGTGCGCC 80
QY 61 CGACGCCCGGGTTTGCOCGGCGGCTCAGTACGACGAGCGCGCGAGAGGGAGTG 120
Db 81 CGACGCCCGGGTTTGCOCGGCGGCTCAGTACGACGAGCGCGCGAGAGGGAGTG 140
QY 121 GCCAACTCCATCACTAGGGTTCTT 145
Db 141 GCCAACTCCATCACTAGGGTTCTT 165
RESULT 8
US-08-254-358-1
; Sequence 1, Application US/08254358
; Patent No. 5658785
; GENERAL INFORMATION:
; APPLICANT: Johnson, Philip R.
; TITLE OF INVENTION: Adeno-Associated Virus Materials and
; NUMBER OF INVENTION: Methods
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US/08/475,391
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/254,358
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 578621land, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31975
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4680 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
```

US-08-475-391-1

Query Match 100.0%; Score 145; DB 1; Length 4680;
Best Local Similarity 100.0%; Fred. No. 5.8e-31;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 60
DB 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 60
QY 61 CGACGCCCGGGCTTTGCCCGCGCGCGCTCACTGAGCGCGCGCGCGCGAGGAGGAGTG 120
DB 61 CGACGCCCGGGCTTTGCCCGCGCGCGCTCACTGAGCGCGCGCGCGCGAGGAGGAGTG 120
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
DB 121 GCCAACTCCATCACTAGGGGTTCT 145

RESULT 10

US-08-709-609-1
Sequence 1, Application US/08709609
Patent No. 5858775
GENERAL INFORMATION:
APPLICANT: Johnson, Philip R.
TITLE OF INVENTION: Adeno-Associated Virus Materials and
METHODS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/709,609
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5858775and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31975
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4680 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

US-08-709-609-1

Query Match 100.0%; Score 145; DB 2; Length 4680;
Best Local Similarity 100.0%; Fred. No. 5.8e-31;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 60
DB 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 60
QY 61 CGACGCCCGGGCTTTGCCCGCGCGCGCTCACTGAGCGCGCGCGCGAGGAGGAGTG 120
DB 61 CGACGCCCGGGCTTTGCCCGCGCGCGCTCACTGAGCGCGCGCGCGAGGAGGAGTG 120

QY 121 GCCAACTCCATCACTAGGGGTTCT 145
DB 121 GCCAACTCCATCACTAGGGGTTCT 145

RESULT 11

PCT-US95-07178-1
Sequence 1, Application PC/TUS9507178
GENERAL INFORMATION:
APPLICANT: Johnson, Philip R.
TITLE OF INVENTION: Adeno-Associated Virus Materials and
METHODS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07178
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Noland, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31975
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4680 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

PCT-US95-07178-1
Query Match 100.0%; Score 145; DB 5; Length 4680;
Best Local Similarity 100.0%; Fred. No. 5.8e-31;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 60
DB 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 60
QY 61 CGACGCCCGGGCTTTGCCCGCGCGCGCTCACTGAGCGCGCGCGAGGAGGAGTG 120
DB 61 CGACGCCCGGGCTTTGCCCGCGCGCGCTCACTGAGCGCGCGCGAGGAGGAGTG 120
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
DB 121 GCCAACTCCATCACTAGGGGTTCT 145

RESULT 12

US-09-299-141-4
Sequence 4, Application US/09299141
Patent No. 6461606
GENERAL INFORMATION:
APPLICANT: FLOTTE, TERENCE R.
APPLICANT: SONG, SIHONG
APPLICANT: BYRNE, BARRY J.
APPLICANT: MORGAN, MICHAEL
TITLE OF INVENTION: MATERIALS AND METHODS FOR GENE THERAPY

```

; FILE REFERENCE: 4300.011800
; CURRENT APPLICATION NUMBER: US/09/299,141
; EARLIER FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: 60/083,025
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 5932
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:p43C-AT
US-09-299-141-4

Query Match      100.0%; Score 145; DB 4; Length 5932;
Best Local Similarity 100.0%; Pred. No. 5.9e-31;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCAGTCAGGCGCGGCGGACCAAGGTCGCC 60
Db 18 TTGGCCACTCCCTCTCTGCGCGCTCGCTCAGTCAGGCGCGGCGGACCAAGGTCGCC 77
QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTCAGGCGGCGGCGGACCAAGGAGTG 120
Db 78 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTCAGGCGGCGGCGGACCAAGGAGTG 137
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 138 GCCAACTCCATCACTAGGGGTTCT 162

RESULT 13
US-09-299-141-4/c
; Sequence 4, Application US/09299141
; Patent No. 6461606
; GENERAL INFORMATION:
; APPLICANT: FLOTTE, TERENCE R.
; APPLICANT: BYRNE, BARRY J.
; APPLICANT: MORGAN, MICHAEL
; TITLE OF INVENTION: MATERIALS AND METHODS FOR GENE THERAPY
; FILE REFERENCE: 4300.011800
; CURRENT APPLICATION NUMBER: US/09/299,141
; EARLIER FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: 60/083,025
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 5932
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:PLASMID
US-09-299-141-8

Query Match      100.0%; Score 145; DB 4; Length 6142;
Best Local Similarity 100.0%; Pred. No. 5.9e-31;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCAGTCAGGCGCGGCGGACCAAGGTCGCC 60
Db 18 TTGGCCACTCCCTCTCTGCGCGCTCGCTCAGTCAGGCGCGGCGGACCAAGGTCGCC 77
QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTCAGGCGGCGGCGGACCAAGGAGTG 120
Db 78 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTCAGGCGGCGGCGGACCAAGGAGTG 137
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 138 GCCAACTCCATCACTAGGGGTTCT 162

RESULT 15
US-09-299-141-8/c
; Sequence 8, Application US/09299141
; Patent No. 6461606
; GENERAL INFORMATION:
; APPLICANT: FLOTTE, TERENCE R.
; APPLICANT: BYRNE, BARRY J.
; APPLICANT: MORGAN, MICHAEL
; TITLE OF INVENTION: MATERIALS AND METHODS FOR GENE THERAPY
; FILE REFERENCE: 4300.011800
; CURRENT APPLICATION NUMBER: US/09/299,141
; EARLIER FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: 60/083,025
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 6142
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:PLASMID
US-09-299-141-8

Query Match      100.0%; Score 145; DB 4; Length 6142;
Best Local Similarity 100.0%; Pred. No. 5.9e-31;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCAGTCAGGCGCGGCGGACCAAGGTCGCC 60
Db 18 TTGGCCACTCCCTCTCTGCGCGCTCGCTCAGTCAGGCGCGGCGGACCAAGGTCGCC 77
QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTCAGGCGGCGGCGGACCAAGGAGTG 120
Db 78 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTCAGGCGGCGGCGGACCAAGGAGTG 137
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 138 GCCAACTCCATCACTAGGGGTTCT 162
```

```

; FILE REFERENCE: 4300.011800
; CURRENT APPLICATION NUMBER: US/09/299,141
; EARLIER FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: 60/083,025
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 6142
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:PLASMID
US-09-299-141-8

Query Match      100.0%; Score 145; DB 4; Length 6142;
Best Local Similarity 100.0%; Pred. No. 5.9e-31;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCAGTCAGGCGCGGCGGACCAAGGTCGCC 60
Db 18 TTGGCCACTCCCTCTCTGCGCGCTCGCTCAGTCAGGCGCGGCGGACCAAGGTCGCC 77
QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTCAGGCGGCGGCGGACCAAGGAGTG 120
Db 78 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTCAGGCGGCGGCGGACCAAGGAGTG 137
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 138 GCCAACTCCATCACTAGGGGTTCT 162

RESULT 15
US-09-299-141-8/c
; Sequence 8, Application US/09299141
; Patent No. 6461606
; GENERAL INFORMATION:
; APPLICANT: FLOTTE, TERENCE R.
; APPLICANT: BYRNE, BARRY J.
; APPLICANT: MORGAN, MICHAEL
; TITLE OF INVENTION: MATERIALS AND METHODS FOR GENE THERAPY
; FILE REFERENCE: 4300.011800
; CURRENT APPLICATION NUMBER: US/09/299,141
; EARLIER FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: 60/083,025
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 6142
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:PLASMID
US-09-299-141-8

Query Match      100.0%; Score 145; DB 4; Length 6142;
Best Local Similarity 100.0%; Pred. No. 5.9e-31;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCAGTCAGGCGCGGCGGACCAAGGTCGCC 60
Db 18 TTGGCCACTCCCTCTCTGCGCGCTCGCTCAGTCAGGCGCGGCGGACCAAGGTCGCC 77
QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTCAGGCGGCGGCGGACCAAGGAGTG 120
Db 78 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTCAGGCGGCGGCGGACCAAGGAGTG 137
QY 121 GCCAACTCCATCACTAGGGGTTCT 145
Db 138 GCCAACTCCATCACTAGGGGTTCT 162
```

Qy	1	TTGGCCACTCCCTCTCTGTGCGCTCGCTCGCTCACTGAGGCGGGCGACCAAGGTCGCC	60
Db	3288	TTGGCCACTCCCTCTCTGTGCGCTCGCTCGCTCACTGAGGCGGGCGACCAAGGTCGCC	3229
Qy	61	CGACGCCCGGGCTTTGCCCCGGGGCTCAGTGAGCGAGCGCGCAGAGAGGGAGTG	120
Db	3228	CGACGCCCGGGCTTTGCCCCGGGGCTCAGTGAGCGAGCGCGCAGAGAGGGAGTG	3169
Qy	121	GCCAACTCCATCACTAGGGTTCTT	145
Db	3168	GCCAACTCCATCACTAGGGTTCTT	3144

Search completed: April 1, 2004, 17:04:17
Job time : 65.7593 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 1, 2004, 12:58:58 ; Search time 243.815 Seconds
(without alignments)
2526.461 Million cell updates/sec

Title: US-10-620-039-1

Perfect score: 145
Sequence: 1 TTGGCCACTCCCTCTGCG.....CTCCATCACTAGGGTTCCT 145

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 29Jan04.*

- 1: geneseqn1980s.*
- 2: geneseqn1990s.*
- 3: geneseqn2000s.*
- 4: geneseqn2001as.*
- 5: geneseqn2001bs.*
- 6: geneseqn2002s.*
- 7: geneseqn2003as.*
- 8: geneseqn2003bs.*
- 9: geneseqn2003cs.*
- 10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	145	100.0	145	2	AAQ41448
2	145	100.0	145	2	AAQ41448
3	145	100.0	145	2	AAQ41448
4	145	100.0	145	2	AAQ41448
5	145	100.0	145	7	ACCS8491
6	145	100.0	145	8	ACFS8876
7	145	100.0	145	7	ABX93568
8	145	100.0	165	2	AAQ66769
9	145	100.0	165	2	AAQ49462
10	145	100.0	165	4	AAQ03535
11	145	100.0	165	6	ABSG6886
12	145	100.0	165	7	ABV77279
13	145	100.0	207	6	AAQ44621
14	145	100.0	955	6	AAQ37254
15	145	100.0	955	6	AAQ37254
16	145	100.0	987	6	AAQ37261
17	145	100.0	987	6	AAQ37261
18	145	100.0	4414	6	AAQ37260
19	145	100.0	4414	6	AAQ37260
20	145	100.0	4476	6	AAQ37259
21	145	100.0	4476	6	AAQ37259
22	145	100.0	4498	6	AAQ37258
23	145	100.0	4498	6	AAQ37258

24	145	100.0	4675	4	AAH41481
25	145	100.0	4675	4	AAH41481
26	145	100.0	4675	6	ABA02989
27	145	100.0	4675	6	ABSG6884
28	145	100.0	4675	6	ABSG6884
29	145	100.0	4675	9	ABE76508
30	145	100.0	4679	4	AAI65974
31	145	100.0	4679	4	AAI65974
32	145	100.0	4679	6	ABK89694
33	145	100.0	4679	7	ABV76133
34	145	100.0	4680	2	AAQ09008
35	145	100.0	4680	7	ABX14497
36	145	100.0	4680	8	ADA50070
37	145	100.0	4680	9	ADH81075
38	145	100.0	4683	4	AAQ37257
39	145	100.0	4825	6	AAQ37257
40	145	100.0	4825	6	AAQ37257
41	145	100.0	4848	6	AAQ37263
42	145	100.0	4966	6	AAQ37256
43	145	100.0	4966	6	AAQ37256
44	145	100.0	4990	6	AAQ37262
45	145	100.0	4990	6	AAQ37262

ALIGNMENTS

RESULT 1

AAQ41448

ID AAQ41448 standard; DNA; 145 BP.

XX AAQ41448;

XX AAQ41448;

DT 25-MAR-2003 (revised)

DT 27-AUG-1993 (first entry)

XX DE

XX AAV2 inverted terminal repeat.

XX Adeno-associated virus 2; ITR; site-specific integration; vector;

XX cell-specific; gene therapy; haemoglobinopathies; thalassemia; diabetes;

XX sickle cell anaemia; cancer; parvovirus; B19; ss.

XX Adeno-associated virus 2.

XX Adeno-associated virus 2.

XX repeat_region 1..125

XX Location/Qualifiers

XX /tag= a

XX /note= "inverted terminal repeat forming palindromic

XX hairpin"

XX 42..83

XX /tag= b

XX /note= "Flip orientation"

XX WO9309239-A1.

XX 13-MAY-1993.

XX 06-NOV-1992; 92WO-US009769.

XX 08-NOV-1991; 91US-00789917.

XX (RESE) RESEARCH CORP TECHNOLOGIES INC.

XX Srivastava A;

XX WPI; 1993-167704/20.

XX Adeno-associated virus-2 basal vectors - for gene therapy and treatment

XX of haemoglobinopathies and cancer etc. - has cassette contg. a promoter

XX capable of cell-specific expression, between 2 inverted terminal repeats

XX of the adeno-associated virus 2.

XX Disclosure; Fig 1; 68pp; English.

XX The sequence is one of two inverted terminal repeat sequences, one from
 CC parvovirus B19 and the other from adeno-associated virus 2 (AAV2), used
 CC in construction of an expression vector for site specific integration and
 CC cell specific gene expression. The vector comprises at least one cassette
 CC contg. a promoter capable of effecting cell-specific expression, operably
 CC linked to a hetero- logous gene, and the cassette residing between the
 CC inverted terminal repeats. The vector is safe for use in gene therapy,
 CC partic. in treatment of haemoglobinopathies and a variety of diseases,
 CC e.g. thalassemia, diabetes, sickle cell anaemia, and cancer. See also
 CC AAO41449. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 145 BP; 21 A; 52 C; 49 G; 23 T; 0 U; 0 Other;
 Query Match 100.0%; Score 145; DB 2; Length 145;
 Best Local Similarity 100.0%; Pred. No. 8.1e-30;
 Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGGACCAAGGTGCGCC 60
 DB 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGGACCAAGGTGCGCC 60
 QY 61 CGACGCCCGGGCTTTGCGCGCGGCGCTCACTGAGCGCGGCGGACCAAGGTGCGCC 120
 DB 61 CGACGCCCGGGCTTTGCGCGCGGCGCTCACTGAGCGCGGCGGACCAAGGTGCGCC 120
 QY 121 GCCAACTCCATCACTAGGGGTTCTCT 145
 DB 121 GCCAACTCCATCACTAGGGGTTCTCT 145
 RESULT 2
 AAT03385
 ID AAT03385 standard; cDNA; 145 BP.
 XX
 AC AAT03385;
 AC
 DT 16-OCT-2003 (revised)
 DT 18-APR-1996 (first entry)
 XX
 DE Strict inverted terminal repeat from AAV-2, used in pITRFL.
 DE
 KW inverted terminal repeat; ITR; adenovirus; adeno-associated virus; AAV-2;
 KW replication defective; integration; gene therapy; ds.
 XX
 OS Adeno-associated virus 2.
 XX
 PN WO9523867-A1.
 PD 08-SEP-1995.
 XX
 PF 28-FEB-1995; 95WO-FR000233.
 XX
 PR 03-MAR-1994; 94FR-00002445.
 XX
 PA (RHON) RHONE POULENC RORER SA.
 XX
 PI Denefle P, Latta M, Perricaudet M, Vigne E;
 XX
 DR WPI; 1995-320581/41.
 XX
 PT Recombinant defective adenovirus contg. integratable expression cassette
 PT - for use in gene therapy to express protein, antigen or anti-sense
 PT nucleic acid, also for prodn. of recombinant adeno-associated viruses.
 XX
 PS Example 3; Page 26; 50pp; French.
 XX
 CC Recombinant, non-pathogenic adenovirus which are able to integrate stably
 CC into a host genome are claimed. The viruses pref. contain at least one
 CC inverted terminal repeat (ITR) sequence and in particular two ITRs flank
 CC a heterologous DNA insert. The present sequence is that of the strict ITR
 CC from adeno associated virus AAV-2 (i.e. the ITR sequence without any
 CC deletions or additions). In the construct pITRFL, the beta-galactosidase

CC marker gene is flanked by two strict AAV-2 ITRs. The defective viruses
 CC are useful for stably introducing large fragments of heterologous DNA
 CC making them suitable for gene therapy. (Updated on 16-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 145 BP; 21 A; 52 C; 49 G; 23 T; 0 U; 0 Other;
 Query Match 100.0%; Score 145; DB 2; Length 145;
 Best Local Similarity 100.0%; Pred. No. 8.1e-30;
 Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGGACCAAGGTGCGCC 60
 DB 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGGACCAAGGTGCGCC 60
 QY 61 CGACGCCCGGGCTTTGCGCGCGGCGCTCACTGAGCGCGGCGGACCAAGGTGCGCC 120
 DB 61 CGACGCCCGGGCTTTGCGCGCGGCGCTCACTGAGCGCGGCGGACCAAGGTGCGCC 120
 QY 121 GCCAACTCCATCACTAGGGGTTCTCT 145
 DB 121 GCCAACTCCATCACTAGGGGTTCTCT 145
 RESULT 3
 AAX34295
 ID AAX34295 standard; DNA; 145 BP.
 XX
 AC AAX34295;
 AC
 DT 16-JUL-1999 (first entry)
 XX
 DE Adeno-associated virus inverted terminal repeat sequence.
 XX
 KW Recombinant; chimeric; parvovirus; adeno-associated virus; AAV; vector;
 KW promoter; rep; cap; inverted terminal repeat; ITR; erythroid cell;
 KW integration; gene expression; bone marrow; peripheral blood cell;
 KW endothelial cell; myocardial cell; ss.
 XX
 OS Adeno-associated virus.
 XX
 PN WO9918227-A1.
 PD 15-APR-1999.
 XX
 PF 08-OCT-1998; 98WO-US021202.
 XX
 PR 08-OCT-1997; 97US-0061364P.
 XX
 PA (ADRE-) ADVANCED RES & TECHNOLOGY INST.
 XX
 PI Srivastava A, Ponnazhagan S;
 XX
 DR WPI; 1999-264033/22.
 XX
 PT New recombinant adeno-associated vectors.
 XX
 PS Claim 2; Page 69; 76pp; English.
 XX
 CC The invention relates to new recombinant chimeric parvovirus-adeno-
 CC associated virus (AAV) vectors comprising a promoter e.g. an AAV promoter
 CC (AAX34296) and a selected DNA sequence, especially an AAV rep gene and a
 CC parvovirus B19 cap gene, located between 2 AAV inverted terminal repeats
 CC (ITR) such as the ITR sequence shown here. The system can specifically
 CC target primitive progenitor and differentiated cells of the erythroid
 CC lineage, and can achieve stable integration and expression of transduced
 CC genes. The vectors can be used for the in vitro or in vivo delivery of
 CC genes to cells such as bone marrow cells, peripheral blood cells,
 CC endothelial cells and myocardial cells
 XX
 SQ Sequence 145 BP; 21 A; 52 C; 49 G; 23 T; 0 U; 0 Other;
 Query Match 100.0%; Score 145; DB 2; Length 145;

[illegible]

CC	disease, receptor for low density lipoprotein gene associated with
CC	familial hypercholesterolemia, hypoxanthine-guanine
CC	phosphoribosyltransferase associated with Lesch-Nyhan syndrome,
CC	phenylalanine hydroxylase (PAH) gene associated with phenylketonuria,
CC	cystinylglycine gene associated with muscular dystrophy, and human cystic
CC	fibrosis transmembrane conductance regulator gene associated with cystic
CC	fibrosis. The present nucleic acid sequence represents a human adeno-
CC	associated virus 2 (AAV2) terminal repeat sequence that was used in the
CC	methods of the invention
XX	
SQ	Sequence 145 BP; 21 A; 52 C; 49 G; 23 T; 0 U; 0 Other;
	Query Match 100.0%; Score 145; DB 6; Length 145;
	Best Local Similarity 100.0%; Pred. No. 8.le-30;
	Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 TTGGGCACCTCCTCTCTCGGCGCTCGTTCGTCTACTGAGCGCGGGGCCAACAAAGGTCGCC 60
Dd	1 TTGGGCACCTCCTCTCTCGGCGCTCGTTCGTCTACTGAGCGCGGGGCCAACAAAGGTCGCC 60
QY	61 CGAAGCGCCGGGGTTTGCCCGGGCGCGCTCACTGAGCGAGCGAGCGGCGAGAGGGAGTG 120
Dd	61 CGAAGCGCCGGGGTTTGCCCGGGCGCGCTCACTGAGCGAGCGAGCGGCGAGAGGGAGTG 120
QY	121 GCACAATCATCACTAGGGGTTCCT 145
Dd	121 GCACAATCATCACTAGGGGTTCCT 145

RESULT 5
 ACC58491/c
 ID ACC58491 standard; DNA; 145 BP.
 XX AC
 AC ACC58491;
 XX AC
 XX AC
 DT 26-AUG-2003 (first entry)
 XX AC
 DE Adeno associated virus inverted terminal repeat region.
 XX AC
 XX AC
 KW AAV; inverted terminal repeat; insect; gene therapy; vector; ss.
 XX AC
 XX AC
 OS Adeno associated virus.
 XX AC
 XX AC
 FH Key Location/Qualifiers
 FT stem_loop 21..145
 FT /*tag= a
 XX AC
 FN WO2003042361-A2.
 XX AC
 XX AC
 PD 22-MAY-2003.
 XX AC
 XX AC
 PF 08-NOV-2002; 2002WO-US035829.
 XX AC
 XX AC
 PR 09-NOV-2001; 2001US-00986618.
 PR 13-AUG-2002; 2002US-00216870.
 XX AC
 XX AC
 FA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX AC
 XX AC
 PI Kotin RM, Urabe M, Ding C;
 XX AC
 XX AC
 DR WPI; 2003-449568/42.
 XX AC
 XX AC
 PT Producing an adeno-associated virus (AAV) in an insect cell, e.g. for
 PT gene therapy, comprises introducing an insect cell-compatible vector into
 PT an insect cell and maintaining the insect cell under conditions that
 PT produce AAV.
 XX AC
 XX AC
 PS Disclosure; Fig 5A; 84pp; English.
 XX AC
 XX AC
 CC The present sequence is that of an adeno associated virus (AAV)
 CC palindromic inverted terminal repeat (ITR) region. An AAV ITR is used in
 CC novel insect cell-compatible vectors of the invention that are designed
 CC for the production of AAV in insect cells. The vectors comprise an AAV

FT	repeat_unit	1..20	
FT		/*tag= b	
FT		/rpt_type= INVERTED	
FT		/label= D_repeat	
FT	misc_structure	21..145	
FT		/*tag= 1	
FT		/label= T-shaped	
FT		/note= "The naturally occurring 145 bp ITR sequence can form a T-shaped structure for DNA replication when single stranded"	
FT	repeat_unit	21..62	
FT		/*tag= c	
FT		/rpt_type= INVERTED	
FT		/label= A' repeat	
FT		/note= "Forms the stem of a T-shaped structure when base paired with repeat A"	
FT	stem_loop	61..82	
FT		/*tag= j	
FT		/note= "Part of T-shaped structure, formed by base pairing between the C' and C repeats"	
FT	repeat_unit	63..70	
FT		/*tag= d	
FT		/rpt_type= INVERTED	
FT		/label= C' repeat	
FT	repeat_unit	74..81	
FT		/*tag= e	
FT		/rpt_type= INVERTED	
FT		/label= C_repeat	
FT	stem_loop	84..104	
FT		/*tag= k	
FT		/note= "Part of T-shaped structure, formed by base pairing between the B' and B repeats"	
FT	repeat_unit	85..92	
FT		/*tag= f	
FT		/rpt_type= INVERTED	
FT		/label= B_repeat	
FT	repeat_unit	96..103	
FT		/*tag= g	
FT		/rpt_type= INVERTED	
FT		/label= B' repeat	
FT	repeat_unit	104..145	
FT		/*tag= h	
FT		/rpt_type= INVERTED	
FT		/label= A_repeat	
FT		/note= "Forms the stem of a T-shaped structure when base paired with A' repeat"	
FT	repeat_unit	146..165	
FT		/*tag= i	
FT		/label= D'	
FT		/note= "Additional D' sequence"	
XX			
PN	WO9636364-AL.		
XX			
XX	21-NOV-1996.		
XX			
PP	14-MAY-1996;	96WO-US006786.	
XX			
PR	15-MAY-1995;	95US-00440738.	
XX			
PA	(SAMU/) SAMULSKI R J.		
PA	(XIAO/) XIAO X.		
XX			
PI	Samulski RJ, Xiao X;		
XX			
XX	WPI; 1997-042643/04.		
DR			
PT	Double-D sequence directs adeno-associated virus integration into host genome - used in gene therapy, maintains full length coding sequence of therapeutic gene.		
PT			
XX			
XX	Claim 1; Fig 9; 55pp; English.		
CC	The adeno associated virus (AAV) has a 145 bp inverted terminal repeat		

PT Composition for transferring recombinant DNAs and to express a
 PT polypeptide in a host cell, comprises two recombinant adeno-associated
 PT viruses.
 XX
 PS Disclosure; Page 139; 144pp; English.
 XX
 CC The patent discloses a composition comprising at least two recombinant
 CC adeno-associated viruses (rAAVs). Each rAAV contains a recombinant DNA
 CC molecule comprising three DNA segments linked together. The first and the
 CC third DNA segments comprise 5' and 3' inverted terminal repeats (ITRs)
 CC respectively from the circular intermediate of AAV. The second DNA
 CC segment in each virus is different and does not comprise AAV sequence. It
 CC preferably comprises sequences encoding a therapeutically effective
 CC polypeptide such as the cystic fibrosis transmembrane conductance
 CC receptor gene (CFTR), the erythropoietin (epo) gene, the tyrosine
 CC hyaloxylase gene (Parkinson's syndrome) or the glucocere- brosidase gene
 CC (Gaucher's disease). The circular intermediate of AAV imparts increased
 CC episomal stability and persistence of the vector in the host cell.
 CC Compositions comprising rAAV sequences are useful for transferring
 CC recombinant DNAs to a host cell and express a polypeptide in a host cell.
 CC The recombinant vector is useful in medical therapy, which includes
 CC treatment or prophylaxis of blood disorders (e.g. sickle cell anaemia,
 CC thalassaemia, haemophilia), neurological disorders, such as Alzheimer's
 CC disease, Parkinson's disease, muscle disorders involving skeletal
 CC cardiac or smooth muscle. AAV vector is used as a delivery vehicle for
 CC gene therapy. The present sequence is the ITR DD DNA sequence referred to
 CC as "double sequence". This sequence is disclosed in U.S. Patent No.
 CC 5,478,745. This sequence is not used in the present invention
 XX
 SQ Sequence 165 BP; 27 A; 56 C; 56 G; 26 T; 0 U; 0 Other;
 Query Match 100.0%; Score 145; DB 4; Length 165;
 Best Local Similarity 100.0%; Pred. No. 8.2e-30;
 Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTGGCACTCCCTCTCTCGGGCTCGCTCGCTCACTAGGCGGGCGACCAAGGTCGCC 60
 DB 21 TTGGCACTCCCTCTCTCGGGCTCGCTCGCTCACTAGGCGGGCGACCAAGGTCGCC 80
 QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTACGAGCGAGCGGCGAGAGGGAGTG 120
 DB 81 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTACGAGCGAGCGGCGAGAGGGAGTG 140
 QY 121 GCCAACTCCATCCTAGGGGTTTCCT 145
 DB 141 GCCAACTCCATCCTAGGGGTTTCCT 165
 RESULT 11
 ABS69886
 ID ABS69886 standard; DNA; 165 BP.
 XX
 AC ABS69886;
 XX
 DT 21-NOV-2002 (first entry)
 XX
 DE Human adeno-associated virus 2 terminal repeat DD sequence.
 XX
 KW Vector; adenovirus; adeno-associated; adenosine deaminase gene; receptor;
 KW adenosine deaminase deficiency; severe combined immune deficiency; PAH;
 KW beta-chain; haemoglobin gene; beta-thalassaemia; sickle cell disease;
 KW low density lipoprotein gene; familial hypercholesterolaemia;
 KW hypoxanthine-guanine phosphoribosyltransferase; Lesch-Nyhan syndrome;
 KW phenylalanine hydroxylase gene; gene therapy; phenylketonuria;
 KW dystrophin gene; muscular dystrophy; cystic fibrosis; immunostimulant;
 KW human cystic fibrosis transmembrane conductance regulator gene;
 KW antianemic; antilipemic; nootropic; cytostatic; dermatological;
 KW human adeno-associated virus 2; AAV2; terminal repeat; ds.
 XX
 OS Homo sapiens.
 XX
 PN US2002102731-A1.
 XX

PD 01-AUG-2002.
 XX
 PF 12-FEB-2001; 2001US-00782378.
 XX
 PR 02-OCT-2000; 2000US-0237747P.
 XX
 PA (UINY) UNIV NEW YORK STATE RES FOUND.
 XX
 PI Hearing P, Bahou WF, Sandalon Z, Gnatenko DV;
 XX
 XX WPI; 2002-690619/74.
 DR
 XX Producing vector, by introducing vector having nucleotide sequence,
 PT adenovirus inverted terminal repeats and packaging sequence, and adeno-
 PT associated virus terminal repeat, into cell, and culturing cell.
 XX
 PS Disclosure; Page 8; 191pp; English.
 XX
 CC The present invention relates to a new method of producing a vector. The
 CC method involves introducing recombinant vector having nucleotide sequence
 CC (NS) having 5' and 3' end, left and right inverted terminal repeats of
 CC adenovirus flanking NS, adenovirus packaging sequence linked to inverted
 CC terminal repeat, and adeno-associated virus terminal repeat linked to 3'
 CC end of NS, into cell expressing adenovirus early gene lacking from vector
 CC and culturing cell to produce another vector. The method is useful for
 CC generating vectors, especially mAd vectors. The method is useful in
 CC transferring nucleotide sequences of interest into a cell, for gene
 CC transfer applications (e.g. gene therapy) in vitro, ex vivo and in vivo.
 CC The nucleotide sequences are useful for treating diseases associated with
 CC it, i.e. adenosine deaminase gene associated with adenosine deaminase
 CC deficiency with severe combined immune deficiency, beta-chain of
 CC haemoglobin gene associated with beta-thalassaemia and sickle cell
 CC disease, receptor for low density lipoprotein gene associated with
 CC familial hypercholesterolaemia, hypoxanthine-guanine
 CC phosphoribosyltransferase associated with Lesch-Nyhan syndrome,
 CC phenylalanine hydroxylase (PAH) gene associated with phenylketonuria,
 CC dystrophin gene associated with muscular dystrophy, and human cystic
 CC fibrosis transmembrane conductance regulator gene associated with cystic
 CC fibrosis. The present nucleic acid sequence represents a human adeno-
 CC associated virus 2 (AAV2) terminal repeat sequence that was used in the
 CC methods of the invention
 XX
 SQ Sequence 165 BP; 24 A; 59 C; 53 G; 29 T; 0 U; 0 Other;
 Query Match 100.0%; Score 145; DB 6; Length 165;
 Best Local Similarity 100.0%; Pred. No. 8.2e-30;
 Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTGGCACTCCCTCTCTCGGGCTCGCTCGCTCACTAGGCGGGCGACCAAGGTCGCC 60
 DB 21 TTGGCACTCCCTCTCTCGGGCTCGCTCGCTCACTAGGCGGGCGACCAAGGTCGCC 80
 QY 61 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTACGAGCGAGCGGCGAGAGGGAGTG 120
 DB 81 CGACGCCCGGGCTTTGCCCGGGCGGCTCAGTACGAGCGAGCGGCGAGAGGGAGTG 140
 QY 121 GCCAACTCCATCCTAGGGGTTTCCT 145
 DB 141 GCCAACTCCATCCTAGGGGTTTCCT 165
 RESULT 12
 ABV77279/c
 ID ABV77279 standard; DNA; 165 BP.
 XX
 AC ABV77279;
 XX
 DT 28-MAR-2003 (first entry)
 XX
 DE Nucleotide sequence of pTRT, a circular adeno-associated virus.
 XX
 KW Circular adeno-associated virus; cAAV; replication; hairpin;
 KW gene therapy; pTRT; ss.
 KW

XX Adeno associated virus.
 XX WO200297056-A2.
 XX 05-DEC-2002.
 XX 31-MAY-2002; 2002WO-US017324.
 XX 31-MAY-2001; 2001US-0294797P.
 XX 07-AUG-2001; 2001US-0313007P.
 XX (UYRQ) UNIV ROCKEFELLER.
 XX Kaplitt MG, Mousseatov S;
 XX WPI; 2003-103706/09.
 XX Production of defective viral vectors for gene therapy that are
 XX completely free of helper viral vectors and helper viruses.
 XX Example 4; Page 53; 69pp; English.
 XX The specification describes a nucleotide sequence capable of directing
 XX circular adeno-associated virus (cAAV) replication. This nucleotide
 XX sequence comprises a loop sequence (TGCCAA) flanked on the 5' and 3'
 XX sides by complementary sequences, where a hairpin structure is formed
 XX between the complementary sequences. The cAAVs are useful in gene
 XX therapy, e.g. to treat an acute medical condition a nucleic acid encoding
 XX a therapeutic protein is inserted into the cAAV. The present sequence
 XX represents pTRT, a cAAV that contains a wild-type circularization point
 XX (the TRT domain), consisting of a single ITR flanked by two D-sequences
 XX
 XX Sequence 165 BP; 26 A; 56 C; 56 G; 27 T; 0 U; 0 Other;
 XX
 XX Query Match 100.0%; Score 145; DB 7; Length 165;
 XX Best Local Similarity 100.0%; Pred. No. 8.2e-30;
 XX Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGCGCAAGGTCGCC 60
 XX Db 145 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGCGGCGCAAGGTCGCC 86
 XX
 XX QY 61 CGACGCGCGCGGCTTTGCGCGCGGCGCTCGCTGAGCGAGCGCGCGAGAGGAGTG 120
 XX Db 85 CGACGCGCGGCTTTGCGCGCGGCGCTCGCTGAGCGAGCGCGCGAGAGGAGTG 26
 XX
 XX QY 121 GCCAACTCCATCACTAGGGGTTCT 145
 XX Db 25 GCCAACTCCATCACTAGGGGTTCT 1
 XX
 XX RESULT 13
 XX AAD44621
 XX ID AAD44621 standard; DNA; 207 BP.
 XX AC AAD44621;
 XX XX
 XX DT 13-DEC-2002 (first entry)
 XX XX
 XX DE 165 bp enzyme attachment site (EAS) DNA.
 XX XX
 XX KW Prokaryotic library; candidate protein; nucleic acid modification; NAM;
 XX KW enzyme attachment sequence; EAS; clinical pharmacology; chemical sensor;
 XX KW enzymology; cosmetic research; toxic; environmental safety assessment;
 XX KW nutrient biology; enzyme attachment site; EAS; ds.
 XX XX
 XX OS Unidentified.
 XX XX
 XX PN WO200266653-A2.
 XX PD 29-AUG-2002.
 XX XX

PF 14-DEC-2001; 2001WO-US049058.
 XX
 PR 14-DEC-2000; 2000US-0256163P.
 XX
 PA (XENC-) KENCOR INC.
 XX
 XX Li M, Liu Y;
 XX
 XX WPI; 2002-667068/71.
 DR
 XX
 XX New library of prokaryotic pET-24a expression vectors, host cells or
 PT nucleic acid/protein conjugates, useful for screening candidate proteins
 PT and their nucleic acids or modification enzymes for pharmacogenetic
 PT analysis.
 XX
 XX Disclosure; Fig 50B; 127pp; English.
 PS
 XX
 XX The invention relates to methods and compositions for the construction of
 CC prokaryotic libraries expressing candidate proteins and the use of these
 CC libraries to identify candidate proteins and the nucleic acids encoding
 CC them. The invention provides a library of prokaryotic pET-24a vectors
 CC comprising a fusion nucleic acid consisting of a nucleic acid encoding a
 CC nucleic acid modification (NAM) enzyme or a candidate protein, or a
 CC nucleic acid having a T7 promoter operably linked to the NAM enzyme or
 CC the candidate protein, and an enzyme attachment sequence (EAS) recognised
 CC by the NAM enzyme. The library is used for identifying candidate proteins
 CC and nucleic acids encoding these proteins, in screening for NAM enzymes
 CC with decreased toxicity for the host cells, or in identifying novel or
 CC improved EASs, which may be used for understanding cellular processes or
 CC any subsequent therapeutic or toxic activities. The nucleic acid/protein
 CC (NAP) conjugates are useful in diagnostic assays and in research
 CC including clinical pharmacology, functional genomics, pharmacogenomics,
 CC agricultural chemicals, environmental safety assessment, chemical sensor,
 CC nutrient biology, cosmetic research or enzymology. These may also be used
 CC in in vitro screening techniques and in assays with target molecules. The
 CC present sequence is an enzyme attachment site (EAS) DNA used in the
 CC invention
 XX
 XX Sequence 207 BP; 36 A; 68 C; 68 G; 35 T; 0 U; 0 Other;
 XX
 XX Query Match 100.0%; Score 145; DB 6; Length 207;
 XX Best Local Similarity 100.0%; Pred. No. 8.3e-30;
 XX Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGCAAGGTCGCC 60
 XX Db 42 TTGGCCACTCCCTCTCTGCGCGCTCGCTCGCTCACTGAGCGCGGCGCAAGGTCGCC 101
 XX
 XX QY 61 CGACGCGCGGCTTTGCGCGGCGCTCGCTGAGCGAGCGCGCGAGAGGAGTG 120
 XX Db 102 CGACGCGCGGCTTTGCGCGGCGCTCGCTGAGCGAGCGCGCGAGAGGAGTG 161
 XX
 XX QY 121 GCCAACTCCATCACTAGGGGTTCT 145
 XX Db 162 GCCAACTCCATCACTAGGGGTTCT 186
 XX
 XX RESULT 14
 XX AAD37254
 XX ID AAD37254 standard; DNA; 955 BP.
 XX AC AAD37254;
 XX XX
 XX DT 21-AUG-2002 (first entry)
 XX XX
 XX DE Adeno-associated virus (AAV) vector plasmid #1.
 XX KW Human; dystrophin minigene; muscular; gene therapy; utrophin; spectrin;
 XX KW adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
 XX KW Becker muscular dystrophy; ds.
 XX XX
 XX OS Homo sapiens.
 XX Unidentified.

```
OS Chimeric.
XX
XX WO200183695-A2.
XX
XX PD 08-NOV-2001.
XX
XX PF 27-APR-2001; 2001WO-US013677.
XX
XX PR 28-APR-2000; 2000US-0200777P.
XX
XX PA (XIAO/) XIAO X.
XX
XX PI Xiao X;
XX
XX DR WPI; 2002-049342/06.
XX
XX PT New dystrophin minigene for treating Duchenne or Becker muscular
XX dystrophy comprises an N-terminal domain or modified N-terminal domain,
XX rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
XX gene.
XX
XX PS Example 1; Page 57; 71pp; English.
XX
XX CC The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified N-
XX terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XX domains and cysteine-rich domains of dystrophin or utrophin genes. The
XX invention also relates to a recombinant adeno-associated virus (AAV)
XX comprising dystrophin minigene operably linked to an expression control
XX element. The dystrophin minigene in operable linkage with an expression
XX control element, in a recombinant adeno-associated virus or retrovirus is
XX useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
XX dystrophy (BMD) in a mammalian subject. The present sequence is AAV
XX vector plasmid containing human dystrophin minigenes, a muscle creatine
XX kinase (MCK) promoter and a small polyA signal sequence
XX
XX SQ Sequence 955 BP; 177 A; 307 C; 296 G; 175 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 145; DB 6; Length 955;
XX Best Local Similarity 100.0%; Pred. No. 9.1e-30;
XX Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTGAGCGCGCGCGCAAGGTCGCC 60
XX
XX Db 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTGAGCGCGCGCGCAAGGTCGCC 60
XX
XX QY 61 CGACGCGCGCGCTTTGCCGCGCGCTCGCTCACTGAGCGCGCGCGCGAGGAGTG 120
XX
XX Db 61 CGACGCGCGCGCTTTGCCGCGCGCTCGCTCACTGAGCGCGCGCGCGAGGAGTG 120
XX
XX QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
XX
XX Db 121 GCCAACTCCATCACTAGGGTTTCCT 145
XX
XX RESULT 15
XX AAD37254/c
XX ID AAD37254 standard; DNA; 955 BP.
XX
XX AC AAD37254;
XX
XX XX
XX DT 21-AUG-2002 (first entry)
XX
XX DE Adeno-associated virus (AAV) vector plasmid #1.
XX
XX KW Human; dystrophin minigene; muscular; Gene therapy; utrophin; spectrin;
XX adeno-associated virus; AAV; Duchenne muscular dystrophy; DMD; BMD;
XX Becker muscular dystrophy; ds.
XX
XX OS Homo sapiens.
XX OS Unidentified.
XX OS Chimeric.
XX
```

```
PN WO200183695-A2.
XX
XX PD 08-NOV-2001.
XX
XX PF 27-APR-2001; 2001WO-US013677.
XX
XX PR 28-APR-2000; 2000US-0200777P.
XX
XX PA (XIAO/) XIAO X.
XX
XX PI Xiao X;
XX
XX DR WPI; 2002-049342/06.
XX
XX PT New dystrophin minigene for treating Duchenne or Becker muscular
XX dystrophy comprises an N-terminal domain or modified N-terminal domain,
XX rod repeats, H1 and H4 domains and a cysteine rich domain of a dystrophin
XX gene.
XX
XX PS Example 1; Page 57; 71pp; English.
XX
XX CC The present invention relates to an isolated nucleotide sequence encoding
XX a dystrophin minigene. The minigene comprises N-terminal or modified N-
XX terminal domains, rod repeats, utrophin or a spectrin gene, H1 and H4
XX domains and cysteine-rich domains of dystrophin or utrophin genes. The
XX invention also relates to a recombinant adeno-associated virus (AAV)
XX comprising dystrophin minigene operably linked to an expression control
XX element. The dystrophin minigene in operable linkage with an expression
XX control element, in a recombinant adeno-associated virus or retrovirus is
XX useful for treating Duchenne muscular dystrophy (DMD) and Becker muscular
XX dystrophy (BMD) in a mammalian subject. The present sequence is AAV
XX vector plasmid containing human dystrophin minigenes, a muscle creatine
XX kinase (MCK) promoter and a small polyA signal sequence
XX
XX SQ Sequence 955 BP; 177 A; 307 C; 296 G; 175 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 145; DB 6; Length 955;
XX Best Local Similarity 100.0%; Pred. No. 9.1e-30;
XX Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTGAGCGCGCGCGCAAGGTCGCC 60
XX
XX Db 955 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTGAGCGCGCGCGCAAGGTCGCC 896
XX
XX QY 61 CGACGCGCGCGCTTTGCCGCGCGCTCGCTCACTGAGCGCGCGCGCGAGGAGTG 120
XX
XX Db 895 CGACGCGCGCGCTTTGCCGCGCGCTCGCTCACTGAGCGCGCGCGCGAGGAGTG 836
XX
XX QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
XX
XX Db 835 GCCAACTCCATCACTAGGGTTTCCT 811
XX
XX Search completed: April 1, 2004, 15:07:23
XX Job time : 249.815 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 1, 2004, 14:05:53 ; Search time 1825.39 Seconds
(without alignments)
3442.959 Million cell updates/sec

Title: US-10-620-039-1
Perfect score: 145
Sequence: 1 TTGGCCACTCCCTCTGCG.....CTCCATCACTAGGGTTCCT 145

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_man.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	145	100.0	145	6	A46401
2	145	100.0	145	14	AA2LTR1
3	145	100.0	145	14	AA2REPORI
4	145	100.0	165	6	AR034135
5	145	100.0	165	6	I16806
6	145	100.0	165	6	AR223306
7	145	100.0	165	6	AX106702
8	145	100.0	165	6	BD18219
9	145	100.0	207	6	AX703496
10	145	100.0	2116	14	AA2LBTFT
11	145	100.0	4675	6	AX135805
12	145	100.0	4675	6	AX286292
13	145	100.0	4675	6	AX753252
14	145	100.0	4675	6	BD094552
15	145	100.0	4675	14	AA2CG
16	145	100.0	4679	6	AX282480
17	145	100.0	4679	14	AF043303
18	145	100.0	4680	6	AR028767
19	145	100.0	4680	6	I62303
20	145	100.0	4681	6	BD242774
21	145	100.0	4683	6	BD242775
22	145	100.0	4683	14	AF028704
23	145	100.0	5932	6	AR235457
24	145	100.0	5932	6	AR235457
25	145	100.0	6142	6	AR235461
26	145	100.0	6142	6	AR235461
27	145	100.0	6565	6	AR235454
28	145	100.0	6565	6	AR235454
29	145	100.0	6714	6	AR235459
30	145	100.0	6714	6	AR235459
31	145	100.0	6924	6	AR235462
32	145	100.0	6924	6	AR235462
33	145	100.0	6924	6	AR235463
34	145	100.0	6924	6	AR235463
35	145	100.0	6924	6	AR235464
36	145	100.0	6924	6	AR235464
37	145	100.0	6981	6	AR235460
38	145	100.0	6981	6	AR235460
39	145	100.0	7054	6	AR235456
40	145	100.0	7054	6	AR235456
41	145	100.0	7405	6	AR235455
42	145	100.0	7405	6	AR235455
43	145	100.0	7492	6	AR235458
44	145	100.0	7492	6	AR235458
45	145	100.0	8698	6	AR222044

ALIGNMENTS

RESULT 1	A46401	A46401	Sequence 4 from Patent WO9523867.	145 bp	DNA	linear	PAT 07-MAR-1997
LOCUS	A46401	A46401	Sequence 4 from Patent WO9523867.				
DEFINITION	A46401	A46401	Sequence 4 from Patent WO9523867.				
ACCESSION	A46401	A46401	Sequence 4 from Patent WO9523867.				
VERSION	A46401.1	GI:2300602					
KEYWORDS	unidentified						
SOURCE	unidentified						
ORGANISM	unclassified.						
REFERENCE	1 (bases 1 to 145)						
AUTHORS	Densfle,P., Latta,M., Perricaudet,M. and Vigne,E.						
TITLE	INTEGRATIVE RECOMBINANT ADENOVIRUSES, PREPARATION THEREOF AND THERAPEUTICAL USES THEREOF						
JOURNAL	Patent: WO 9523867-A 4 08-SEP-1995;						

Pred. No. is the number of results predicted by chance to have a

COMMENT
RHONE POULENC ROBER SA (FR)
Other publication AU 1852695 950918
Other publication FR 2716893 950908.
FEATURES
source
1. .145
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 100.0%; Score 145; DB 6; Length 145;
Best Local Similarity 100.0%; Pred. No. 1.5e-23;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTAGGCGCGGCGGACCAAGGTGCC 60
Db 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTAGGCGCGGCGGACCAAGGTGCC 60
Qy 61 CGACGCCCGGGTTTGGCCGGCGGCTTCAGTGAGCGAGCGGCGGCGGAGGAGTG 120
Db 61 CGACGCCCGGGTTTGGCCGGCGGCTTCAGTGAGCGAGCGGCGGCGGAGGAGTG 120
Qy 121 GCCAACTCCATCACTAGGGTTTCCT 145
Db 121 GCCAACTCCATCACTAGGGTTTCCT 145

RESULT 2

AA2LTR1 145 bp DNA linear VRL 27-APR-1993
LOCUS
DEFINITION Adeno-associated virus 2 left terminal sequence.
ACCESSION K01624
VERSION K01624.1 GI:209623
KEYWORDS replication; terminal repeat.
SEGMENT 1 of 2
SOURCE Adeno-associated virus 2H
ORGANISM Adeno-associated virus 2H
VIRUSES; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
REFERENCE 1 (bases 1 to 145)
AUTHORS Lusby, E., Fife, K.H. and Berns, K.I.
TITLE Nucleotide sequence of the inverted terminal repetition in adeno-associated virus DNA
J. Virol. 34 (2), 402-409 (1980)
MEDLINE 80185149
PUBMED 6246271
REFERENCE 2 (bases 1 to 145)
AUTHORS Lefebvre, R.B., Riva, S. and Berns, K.I.
TITLE Conformation takes precedence over sequence in adeno-associated virus DNA replication
Mol. Cell. Biol. 4 (7), 1416-1419 (1984)
MEDLINE 85061247
PUBMED 6504049
COMMENT Original source text: Adeno-associated virus 2H DNA, (clone pSM620 [2]), from KB or HeLa cells.
Both [1] and [2] present the opposite strand from the one presented here. The focus of both papers is the method of replication of the virus. [1] notes that the initial tt is present only 30% of the time; it is shortened to t in 50% of the population and missing altogether in 15% of the population. There is further sequence heterogeneity which can be explained by assuming that the terminal 125 bases, which form an imperfect palindrome, are replaced by their inverted complement during replication. [2] found that deletion of the 9 terminal bases on the right and the 113 terminal bases on the left of AAV 2 genome did not stop DNA replication. Further deletion of an 11-base symmetrical sequence (bases 89 to 99) in the right terminal repetition inhibits DNA replication. Substitution of either an 8-base (cagatctg) or 12-base (cgggtagcgcg) symmetrical sequence unrelated to the original 11-base sequence restores DNA replication. All of this can be explained by assuming that the 125 base palindrome mentioned above form a t-shaped secondary structure which provides a primer for DNA polymerase during replication.

FEATURES

Location/Qualifiers

source
1. .145
/organism="Adeno-associated virus 2H"
/mol_type="genomic DNA"
/db_xref="taxon:10805"

ORIGIN

2 bases upsteam of HaeIII site.
Query Match 100.0%; Score 145; DB 14; Length 145;
Best Local Similarity 100.0%; Pred. No. 1.5e-23;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTAGGCGCGGCGGACCAAGGTGCC 60
Db 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTAGGCGCGGCGGACCAAGGTGCC 60
Qy 61 CGACGCCCGGGTTTGGCCGGCGGCTTCAGTGAGCGAGCGGCGGCGGAGGAGTG 120
Db 61 CGACGCCCGGGTTTGGCCGGCGGCTTCAGTGAGCGAGCGGCGGCGGAGGAGTG 120
Qy 121 GCCAACTCCATCACTAGGGTTTCCT 145
Db 121 GCCAACTCCATCACTAGGGTTTCCT 145

RESULT 3

AA2REPORI 145 bp ss-DNA linear VRL 27-APR-1993
LOCUS
DEFINITION Adeno-associated virus origin of replication (genome 3' terminus).
ACCESSION M10681
VERSION M10681.1 GI:209626
KEYWORDS
SOURCE Adeno-associated virus 2H
ORGANISM Adeno-associated virus 2H
VIRUSES; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
REFERENCE 1 (bases 1 to 145)
AUTHORS Berns, K.I., Hauswirth, W.W., Fife, K.H. and Lusby, E.
TITLE Adeno-associated virus DNA replication
JOURNAL Cold Spring Harb. Symp. Quant. Biol. 43 Pt 2, 781-787 (1979)
MEDLINE 80023388
PUBMED 226321
COMMENT Original source text: Adeno associated virus 2H (AAV2 H) DNA.
FEATURES
source
1. .145
/organism="Adeno-associated virus 2H"
/mol_type="genomic DNA"
/db_xref="taxon:10805"

ORIGIN

Query Match 100.0%; Score 145; DB 14; Length 145;
Best Local Similarity 100.0%; Pred. No. 1.5e-23;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTAGGCGCGGCGGACCAAGGTGCC 60
Db 145 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTAGGCGCGGCGGACCAAGGTGCC 86
Qy 61 CGACGCCCGGGTTTGGCCGGCGGCTTCAGTGAGCGAGCGGCGGCGGAGGAGTG 120
Db 85 CGACGCCCGGGTTTGGCCGGCGGCTTCAGTGAGCGAGCGGCGGCGGAGGAGTG 26
Qy 121 GCCAACTCCATCACTAGGGTTTCCT 145
Db 25 GCCAACTCCATCACTAGGGTTTCCT 1

RESULT 4

AR034135 165 bp DNA linear PAT 29-SEP-1999
LOCUS
DEFINITION Sequence 1 from patent US 5869305.
ACCESSION AR034135
VERSION AR034135.1 GI:5949740
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.

```

REFERENCE 1 (bases 1 to 165)
AUTHORS Samulski, R. Jude, and Xiao, X.
TITLE Recombinant viral vector system
JOURNAL Patent: US 5869305-A 1 09-FEB-1999;
FEATURES Location/Qualifiers
source 1..165
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 145; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.5e-23;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 60
Db 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 80

QY 61 CGAGCGCGCGGCTTTGCCCGGCGCTCACTGAGCGAGCGCGCGAGAGGGAGTG 120
Db 81 CGAGCGCGGCGCTTTGCCCGGCGCTCACTGAGCGAGCGCGCGAGAGGGAGTG 140

QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
Db 141 GCCAACTCCATCACTAGGGTTTCCT 165

RESULT 5
LOCUS I16806 165 bp DNA linear PAT 03-APR-1996
DEFINITION Sequence 1 from patent US 5478745.
ACCESSION I16806
VERSION I16806.1 GI:1251714
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 165)
AUTHORS Samulski, R. J. and Xiao, X.
TITLE Recombinant viral vector system
JOURNAL Patent: US 5478745-A 1 26-DEC-1995;
FEATURES Location/Qualifiers
source 1..165
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 145; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.5e-23;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 60
Db 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 80

QY 61 CGAGCGCGGCTTTGCCCGGCGCTCACTGAGCGAGCGCGCGAGAGGGAGTG 120
Db 81 CGAGCGCGGCTTTGCCCGGCGCTCACTGAGCGAGCGCGCGAGAGGGAGTG 140

QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
Db 141 GCCAACTCCATCACTAGGGTTTCCT 165

REFERENCE 1 (bases 1 to 165)
AUTHORS Samulski, R. J. and Xiao, X.
TITLE Recombinant viral vector system
JOURNAL Patent: US 5869305-A 1 09-FEB-1999;
FEATURES Location/Qualifiers
source 1..165
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 145; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.5e-23;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 60
Db 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 80

QY 61 CGAGCGCGGCTTTGCCCGGCGCTCACTGAGCGAGCGCGCGAGAGGGAGTG 120
Db 81 CGAGCGCGGCTTTGCCCGGCGCTCACTGAGCGAGCGCGCGAGAGGGAGTG 140

QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
Db 141 GCCAACTCCATCACTAGGGTTTCCT 165

RESULT 6
LOCUS AR223306 165 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 7 from patent US 6436392.
ACCESSION AR223306
VERSION AR223306.1 GI:23331457
KEYWORDS
SOURCE Unknown.

```

```

ORGANISM Unknown.
REFERENCE 1 (bases 1 to 165)
AUTHORS Engelhardt, J. F. and Duan, D.
TITLE Adeno-associated virus vectors
JOURNAL Patent: US 6436392-A 7 20-AUG-2002;
FEATURES Location/Qualifiers
source 1..165
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 145; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.5e-23;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 60
Db 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 80

QY 61 CGAGCGCGGCTTTGCCCGGCGCTCACTGAGCGAGCGCGCGAGAGGGAGTG 120
Db 81 CGAGCGCGGCTTTGCCCGGCGCTCACTGAGCGAGCGCGCGAGAGGGAGTG 140

QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
Db 141 GCCAACTCCATCACTAGGGTTTCCT 165

RESULT 7
LOCUS AX106702 165 bp DNA linear PAT 30-APR-2001
DEFINITION Sequence 7 from Patent WO0125465.
ACCESSION AX106702
VERSION AX106702.1 GI:13922363
KEYWORDS
SOURCE unidentified
ORGANISM unidentified.
REFERENCE 1
AUTHORS Engelhardt, J. F., Dongsheng, D. and Ziying, Y.
TITLE Adeno-associated viruses and uses thereof
JOURNAL Patent: WO 0125465-A 7 12-APR-2001;
University of Iowa Research Foundation (US); Engelhardt, John F. (US); Dongsheng, Duan (US); Ziying, Yan (US)
FEATURES Location/Qualifiers
source 1..165
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
/note="SSEQ ID NO:1 of U.S. Patent No. 5,478,745"

ORIGIN
Query Match 100.0%; Score 145; DB 6; Length 165;
Best Local Similarity 100.0%; Pred. No. 1.5e-23;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 60
Db 21 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGCGCGACCAAGGTCGCC 80

QY 61 CGAGCGCGGCTTTGCCCGGCGCTCACTGAGCGAGCGCGCGAGAGGGAGTG 120
Db 81 CGAGCGCGGCTTTGCCCGGCGCTCACTGAGCGAGCGCGCGAGAGGGAGTG 140

QY 121 GCCAACTCCATCACTAGGGTTTCCT 145
Db 141 GCCAACTCCATCACTAGGGTTTCCT 165

RESULT 8
LOCUS BD218219 165 bp DNA linear PAT 17-JUL-2003

```


JOURNAL Patent: WO 0132711-A 5 10-MAY-2001;
THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ARKANSAS (US)
FEATURES
source
1. .4675
/organism="Adeno-associated virus 2"
/mol_type="unassigned DNA"
/db_xref="taxon:10804"
321. .2186
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC39573.1"
/db_xref="GI:14272041"
/db_xref="REVIEW:MEL:CA39573"
/translation="MPGFYELVIVKPSDLQGLHPLGISDSFVNVVAEKELPDPDMD
LNLEAPLVAELQLQDLFLTEWRVSKAFELFFVQFEGESYFNMHLVETGVKS
MVLGRFLSQIRKLIQRIYRGIETPLNFAVTKTRNGAGGKGVDECVIPYLLPK
TQPLEWANTNMEQLSACNLTKRLVAQHLTVSQTEQKKNQNPNSDAPVRS
KTSARYMELVGLVKGIDTQKQWQIDQASYISFNAASRSQIKAAALNAGKIMSL
TKTAPDVLVGQPPVEDISNNRIYKILNGVDPQYAAASVFLGWATKFGKRTIMLFG
PATGKNIAEIAHTVFPYGCNVNTNENPFNDVCDKWIWEEGKMTAKVYESAKA
ILGSKVYDOKCKSSAQIDPTPIVTSNTMCAVIDNSTTPEHQOPLDRMEKPEL
TRRDHFGKVTKEVDFFRNAXDHWVEHEFYVKGAKRPARPSADADISEPKRV
RESVAQTSDAEASINADRYQNKSRHVMNMLPFCQCRNMNNSNICFTHGQK
DCLECFPVSESPVSVKAYQKLCYIHIMNGKVPDCTACTCDLVNVDLDCIFEQ"

ORIGIN
Query Match 100.0%; Score 145; DB 6; Length 4675;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTAGAGCGCGGCGCAAAAGTGC 60
DB 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTAGAGCGCGGCGCAAAAGTGC 60

QY 61 CGACGCCGGGCTTGGCCGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
DB 61 CGACGCCGGGCTTGGCCGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120

QY 121 GCCAACTCCATCACTAGGGGTTCT 145
DB 121 GCCAACTCCATCACTAGGGGTTCT 145

RESULT 12
AX286292
LOCUS 4675 bp DNA linear PAT 21-NOV-2001
DEFINITION Sequence 1 from Patent WO0180840.
ACCESSION AX286292
VERSION AX286292.1 GI:17048540
KEYWORDS Adeno-associated virus 2
SOURCE Adeno-associated virus 2
ORGANISM Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
REFERENCE 1
AUTHORS Raj, K. and Beard, P.M.
TITLE Cytotoxic agents
JOURNAL Patent: WO 0180840-A 1 01-NOV-2001;
BIG INTERNATIONAL LIMITED (GB)
FEATURES
source
1. .4675
/organism="Adeno-associated virus 2"
/mol_type="unassigned DNA"
/db_xref="taxon:10804"

ORIGIN
Query Match 100.0%; Score 145; DB 6; Length 4675;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTAGAGCGCGGCGCAAAAGTGC 60
DB 1 TTGGCCACTCCCTCTCTCGCGCTCGCTCGCTCACTAGAGCGCGGCGCAAAAGTGC 60

QY 61 CGACGCCGGGCTTGGCCGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120
DB 61 CGACGCCGGGCTTGGCCGGCGGCTCAGTGAGCGAGCGCGCGAGAGGGAGTG 120

QY 121 GCCAACTCCATCACTAGGGGTTCT 145
DB 121 GCCAACTCCATCACTAGGGGTTCT 145

RESULT 14
BD094552
LOCUS 4675 bp DNA linear PAT 27-AUG-2002
DEFINITION Method of transferring gene.
ACCESSION BD094552
VERSION BD094552.1 GI:22640140
KEYWORDS WO 0132899-A/1.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 4675)
AUTHORS Ueno, T., Matsumura, H., Tanaka, K., Iwasaki, T., Ueno, M., Fujinaga, K., Asada, K. and Kato, I.
TITLE Method of transferring gene
JOURNAL Patent: WO 0132899-A 1 10-MAY-2001;
TAKARA SHUZO CO LTD, TAKASHI UENO, HAJIME MATSUMURA, KEIJI TANAKA, TOMOKO IWASAKI, MIITSUHIRO UENO, KEI FUJINAGA, KIYOZO ASADA, IKUNOSHIN KATO
OS Adeno-associated virus
PN WO 0132899-A/1
PD 10-MAY-2001
PF 23-OCT-2000 WO 2000JP007373
PR 29-OCT-1999 JP 99P 308839
PI TAKASHI UENO, HAJIME MATSUMURA, KEIJI TANAKA, TOMOKO IWASAKI, PI

MITSUHIRO UENO,	
PI	KEI FUJINAGA, KIYOZO ASADA, IKUNOSHIN KATO
PC	Cl2N15/861, A61K35/12, A61K35/76, A61K48/00, Cl2N5/10, Cl2N7/01 CC
Method of transferring gene	
FH	Key Location/Qualifiers
FT	source 1..4675
FT	/organism='Adeno-associated virus',
Location/Qualifiers	
source	1..4675
/organism='unidentified'	
/mol_type='genomic DNA'	
/db_xref='taxon:32644'	
ORIGIN	
Query Match 100.0%; Score 145; DB 6; Length 4675;	
Best Local Similarity 100.0%; Pred. No. 9.3e-24;	
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGGCCACCAAGGTGCC 60
Db	1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGCGCGGCCACCAAGGTGCC 60
QY	61 CGAGCGCCGGCGCTTTGCGCGCGCGCTCACTGAGCGAGCGCGCAGAGGGAGTG 120
Db	61 CGAGCGCCGGCGCTTTGCGCGCGCGCTCACTGAGCGAGCGCGCAGAGGGAGTG 120
QY	121 GCCAACTCCATCACTAGGGTTCTCT 145
Db	121 GCCAACTCCATCACTAGGGTTCTCT 145
RESULT 15	
AA2CG	4675 bp ss-DNA linear VRL 27-APR-1993
LOCUS	Adeno-associated virus 2, complete genome.
DEFINITION	Adeno-associated virus 2, complete genome.
ACCESSION	J01901 M12405 M12468 M12469
VERSION	J01901.1 GI:209616
KEYWORDS	alternative splicing; complete genome; major coat protein.
SOURCE	Adeno-associated virus 2
ORGANISM	Adeno-associated viruses 2
REFERENCE	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
AUTHORS	Samulski, R.J., Srivastava, A., Berns, K.I. and Muzyczka, N.
TITLE	Rescue of adeno-associated virus from recombinant plasmids: gene correction within the terminal repeats of AAV
JOURNAL	Cell 33 (1), 135-143 (1983)
MEDLINE	84282662
PUBMED	6088052
REFERENCE	2 (bases 1 to 4675)
AUTHORS	Srivastava, A., Lusby, E.W. and Berns, K.I.
TITLE	Nucleotide sequence and organization of the adeno-associated virus 2 genome
JOURNAL	J. Virol. 45 (2), 555-564 (1983)
MEDLINE	83164299
PUBMED	6300419
COMMENT	Original source text: Adeno-associated virus 2 DNA from human Hela cells.
FEATURES	
source	Location/Qualifiers
	1..4675
	/organism='Adeno-associated virus 2'
	/mol_type='Genomic DNA'
	/db_xref='taxon:10804'
	1..145
	/note='5' inverted terminal repeat'
	42..83
	/note='flip oriented DNA'
	287..4447
	/note='major coat protein A mRNA (alt.)'
	join(321..1906,2228..2252)
	/note='major coat protein A' (alt.)'
	/codon_start=1
	/protein_id='AAA42372.1'
	/db_xref='GI:209617'
repeat_region	
misc_feature	
mRNA	
CDS	

```

ORIGIN      5' end of genomic DNA.
Query Match 100.0%; Score 145; DB 14; Length 4675;
Best Local Similarity 100.0%; Pred. No. 9.3e-24;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGGCCGGGGGACCAAGGTCGCC 60
Db 1 TTGGCCACTCCCTCTCTGCGCGCTCGCTCACTGAGGCCGGGGGACCAAGGTCGCC 60

Qy 61 CGAGCCCGGGCTTTGCCCCGGGGCCCTCAGTGAGCGAGCGAGCGAGAGGAGTG 120
Db 61 CGAGCCCGGGCTTTGCCCCGGGGCCCTCAGTGAGCGAGCGAGCGAGAGGAGTG 120

Qy 121 GCCAACTCCATCACTAGGGTTCTT 145
Db 121 GCCAACTCCATCACTAGGGTTCTT 145

```

Search completed: April 1, 2004, 16:04:09
Job time : 1829.39 secs